

Psychiatric care in further trouble

There is a danger that Britain will be frightened by the violence of some psychiatric patients into undoing an enlightened policy on care. More resources are a better answer.

THERE is a comforting ring to the slogan 'care in the community', used in Britain to describe the discharge of psychiatric patients from ageing hospitals into less formal accommodation. Not least, it implies that the institutions built in the nineteenth century and then called bedlams are no longer suited to the needs of patients, let alone seemly to the rest of the population, and that they will be replaced by a different regimen in which patients rub shoulders with people outwardly like themselves and in which they are supported not only by modern drugs but also by the care and compassion of a supposedly sympathetic community. At its best, the concept of community care is a libertarian recognition that psychiatric patients are people, like those who suffer from other diseases, and are deserving of respect even when deprived of liberty.

The movement has prospered marvellously since the early 1960s, notably with the advent of derivatives of the phenothiazines for controlling schizophrenic behaviour, cyclic amines in depressive illness and simple ionic lithium in manic-depressive illness. Palliative though they are, the benefits of these drugs in the treatment of the psychiatrically disabled have amply justified the enthusiasm of three decades ago. Of course, there are snags. Not all patients respond well, or consistently over time. Sometimes there are unpleasant side-effects. Even more serious, patients will not always take the drugs prescribed for them; to do so is to admit disablement. Yet none of that detracts from the value of the drugs, now used to enable one or two per cent of the adult populations of rich societies, who would previously have been ostracized, to live more or less ordinary lives.

Legacies

Other legacies of the 1960s also mark and sometimes mar treatment in the community. The conceit that even madness is only relative led some to conclude that, if the inmates of bedlam and their caretakers are objectively indistinguishable, there can be no basis for locking up members of either group. That has fostered legal and administrative diffidence, which equates with muddle, over the circumstances in which people can be deprived of liberty against their wishes. Another hangover from the 1960s is the expectation that an informal regimen for patients would be cheaper.

Although community care offers the majority of psychiatric patients a taste of life they would previously have been denied by incarceration, the most compassionate expectations of it have not been fulfilled. Many of those decanted

from the old hospitals have ended up not "in the community", but on the edge of it; they are a large proportion of those who live on urban streets in cardboard boxes and who doss down overnight in public parks. Unsurprisingly, they most often congregate in cosmopolitan and affluent cities. Mostly, they are harmless people and so can easily be ignored, as they have been, mostly. But some of them are given to essays in petty crime, shoplifting and even car theft for example. Smart police and prosecution authorities have learnt to recognize these crimes as the appeals they are for anything from a warm bed for the night to a ticket back to the decanting institution. And so, by leniency and indulgence, community care becomes as institutionalized as the old bedlam regimen.

Alarm

The British, although well used to looking the other way, have now been alarmed that psychiatric patients recently released from hospitals have often marked their return to the community by crimes more serious than mere car theft. On one estimate, there have been 30 murders in the past two years by people discharged from psychiatric hospitals within the previous 12 months. Inevitably, many of these cases grab the headlines. The death the other day of a young boy, at the hands of a newly discharged patient, in a manner closely resembling the earlier murder for which the patient was locked up in the first place, has predictably set up a clamour for stricter control. Predictably, the government has responded by promising that those who fail to take the medication prescribed for them will be recalled to the hospitals from which they have been discharged.

The only certainty in that arrangement is that it can be only a partial remedy for a terrible dilemma. Neither the effect of drugs on those whom they benefit nor the assessment of a sick person's condition from day to day can be sure enough for anybody to promise that discharged patients will never, for example, murder other people. At this stage, it is also far from clear to what extent the recall of recalcitrant psychiatric patients will amount to a reversal of the policy of community care. But on the face of things, a policy of 'tightening up' seems to be a recipe for still more tightening up when the next public outcry is provoked. Psychiatrists, not much honoured as things are, are destined to be perpetual scapegoats.

Is there a better way? The keyword is 'community'. Indubitable successes notwithstanding, too many of those discharged from hospitals into community care have only a marginal place in a real community. Often, their illness is as



much to blame as the embarrassment of others; those most in need of help are notoriously the least appealing. But it would be a bad business if the occasional violence of the occasional discharged psychiatric patient were to lead by stealth to an unwinding of the successes of the past several decades. That the amount of violence has become obtrusive is largely a measure of how much psychiatric illness there is. And governments, which may be blamed for many things, cannot be blamed for that. What they will be blamed for is an abandonment of an enlightened framework for the management of the psychiatrically disabled in the face of public clamour, and because the costs (now apparent) are not easily affordable. Their best course, having coined the phrase 'care in the community', is to remind the community of its duties of care and compassion. □

Clean drugs for Italy?

Italy should avoid going to extremes in its attempts to reform the system of drug registration and pricing.

THE callousness of the malign ring of corruption centring on Italy's Ministry of Health (see page 663) defies belief. Former health minister Francesco De Lorenzo has apparently been colluding with drug companies for many years to fix unfair drug prices and to guarantee pharmaceuticals with no proven benefit a place in the national formulary. The shame extends into Italy's universities, where several prominent professors have been accused of accepting bribes to give the right 'advice'. Italians can only be thankful that the process of screening drugs for toxicity took place at a much earlier stage, distant from the influence of politicians.

The new minister for health, Christian Democrat Mariapia Garavaglia, now has the task of creating from scratch 'clean' systems for drug registration and drug pricing. But if she is not careful, overreaction to the corruption may lead to future problems. The recent revelations have led to such a climate of disgust and distrust that Garavaglia has said that she will not create another commission for the fixing of pharmaceutical prices. Instead, she wants to establish a system of free pricing, presumably to be dictated by the open market. The Department of Industry, which by law has to agree all prices, is not of the same mind. Much of the national drug bill is reimbursed by the government, and it believes that the government should therefore have some say in the matter.

Italy should not jump out of the frying pan into the fire. For too long, taxpayers have been asked to foot the bill for overpriced drugs, many of them in any case medically ineffective. They should not now be asked to foot a bill whose size is dictated entirely by the commercial interests of a national industry, which has itself not emerged squeaky clean from the corruption scandal, without any form of government control. The Ministry of Industry has already put together a panel of experts whose job is to decide on criteria for price fixing. The Department of Health should listen to the panel's advice and be prepared to act in the best interests of Italy's health service. □

Irish funding disappears

Ireland needs a champion in the government who will look to the future of the science base.

OVER the past decade, European researchers have come to expect less and less support from their governments. But it is not often that a government agency baldly decides to make savings by stopping all new funding for basic research. This is how Eolas, Ireland's major research funding agency, chose to solve its cash problems earlier this year.

We now hear that a (risibly) small amount of money has been pulled out of the hat to defend basic research (see page 662). But this gesture, hard fought for within the agency, cannot disguise the message behind the original decision: that basic research is held in disregard by Irish policy-makers. Few even bother to pretend that such research is worthwhile. Some go as far as to say that "those university types have had it good for too long".

But worse may be to come. Ireland's Ministry of Employment and Enterprise, to which Eolas is responsible, last month indicated to the Committee of Heads of Irish Universities that it no longer wants to fund basic research in the future. Its budget will be directed entirely to applied research.

As part of a major reorganization, Eolas is being merged this autumn with the Industrial Development Agency (Ireland). Eolas's original optimism that the merger will make no practical difference to its operations is in doubt. No-one from the Eolas board, and no scientist, has been appointed to the Forbairt. Its new director, Eileen O'Mara Walsh, is chairperson of the Great Southern Hotel Group.

So if the Department of Employment and Enterprise no longer takes responsibility for basic research, who will? Although the government agrees that a basic research base should be maintained, in practice a big question mark hangs over the future. The most natural home, the Department of Education, has no plans to create a research programme. It has not been given the resources. (The Department of Health still has, of course, its rather inadequate budget for biomedical research — but even this is having to stretch further to cover new areas such as epidemiology.)

Science has no champion in the Irish government. There is no ministry with special responsibility for science, no research council, not even a major lobby. Ireland has no research institutes analogous to Germany's Max Planck or France's CNRS; almost all basic research is carried out in universities. After years of being ignored, the Committee of Heads of Irish Universities is now demanding more say in policy, but it may be too late. Its main demand is that universities be considered a 'resource for industry'.

Ireland is a poor country. No one blames the government for wanting to improve its industrial competitiveness as soon as possible. And everyone would agree that research for and by industry is fundamental to this aim. But Ireland must also look further than the next few years. Allowing its research base, such as it is, simply to fade away will seed a much greater potential for economic failure in the future. □

Ranchers feel pressure on environment

Washington. The Clinton administration last week began moves to impose higher fees and tighter environmental restrictions on ranchers grazing herds on public land. The step is the first to run counter to the century-old practice of encouraging settlement in the West through land-use subsidies.

The White House removed increases in grazing and mining fees from Clinton's deficit-reduction plan last winter because several Western lawmakers had threatened to withhold their votes. With his economic package now through Congress, Clinton

seems to be trying to dispel the notion that he renounced his environmental principles to buy votes: over the following months the administration will follow up its grazing laws with new regulations on mining, timber and water rights.

Some 31,000 ranchers graze herds on the 272 million acres of federal lands administered by the Bureau of Land Management (BLM) and the Agriculture Department's Forest Service. Under the new plan, they would have to pay \$4.28 per animal unit (five sheep or a cow and calf) per month compared with \$1.86 now. This higher tariff would still be lower than the market rate (\$5 to \$12), but would nonetheless provide the Treasury with an extra \$98 million over the next three years.

"We've found a reasonable balance between the need to sustain the health of rangeland ecosystems and the need to sustain the economic health of rural Western areas," says Bruce Babbitt, Interior Secretary. Babbitt has championed the plan within the administration.

The 'Babbitt plan' would also curtail the use of pesticides, shorten grazing seasons and make renewal of rancher's leases dependent on their environmental record. And it would end a long-standing BLM policy that lets ranchers claim water rights and hold title to range improvements. The plan would also extend membership of BLM's grazing advisory boards to include not only ranchers, but also businesspeople, environmentalists and fish and wildlife experts; the boards help to review applications and set local standards for federal lands.

Conservationists have welcomed Babbitt's plan. They have long complained that the administration has done little to encourage environmental responsibility among commercial users of federal lands. Mismanaged grazing, they say, has caused compacting of soils, erosion and destruction

of plant and animal diversity. Overgrazing in riparian areas, they say, has destroyed fisheries and nesting cover by creating muddy banks and stream runoff, while in desert areas it has threatened species such as the desert tortoise and Sonoran pronghorn by depleting water resources.

"They're running cow factories on land that belongs to all Americans," says Melanie Griffin, the public lands expert for the Sierra Club. "There's no reason why ranchers should be destroying the habitat and getting a free ride."

Ranching interests have denounced the

Babbitt plan. They claim it would deal a mortal economic blow to the livestock industry and to the small rural communities that depend on the ranching community. Industry experts reckon the plans could cause 20,000 ranching operations to close and increase beef prices by as much as a third.

The administration does not need approval from Congress to implement Babbitt's plans. They will probably come into effect some time after 1994, after completing the lengthy federal rule-making process.

Susan Greene.



Babbitt: seeking balance

Push for Gulf syndrome research

Washington. US politicians and veterans' organizations are calling for more research into the effects of battlefield chemicals on health, because thousands of Gulf War veterans have developed a mysterious illness.

Researchers do not know what causes the multiple symptoms (see panel) reported by more than 6,000 US soldiers who served in the Gulf. "Chemical exposure is one of the clearest possibilities", says Terry Jemison, a

spokesman for the US Veterans' Administration. But health officials at the US Department of Defense claim there is no epidemiological evidence for the so-called 'Gulf War syndrome'.



Buyer: diesel a factor

Lack of consensus about 'multi-chemical sensitivity' is nothing new. The US health establishment has long questioned its existence because many feel the diagnostic and therapeutic approaches of clinical ecologists are inadequate.

"Until we develop a rational scientific approach to the problem, thousands of veterans will be caught in crossfire between doctors who consider chemical sensitivity to be real and doctors who don't," says Congressman Steve Buyer (Republican, Indiana). Buyer, who served in the Gulf, blames lead inhaled from a diesel-fuelled tent heater for a series of illnesses.

He and his colleagues in Congress will probably ask for funding for such research when they vote on the defence budget in the autumn. Congress has already authorized the Veterans' Administration and Department of Defense to grant \$500,000 annually for a review of scientific information on the effects of Gulf operations on health.

Claudia Miller, professor of allergy and immunology at the University of Texas

Health Science Center at San Antonio, is doing pioneering research in chemical sensitivity. She has planned a specialized environmental unit to isolate and blindly challenge veterans with chemicals. The direct challenge approach would be the "gold standard", she says, that would settle the debate. Miller estimates that the unit would cost \$1 million to build and \$500,000 a year to run.

Supporters of the veterans' cause applaud plans such as Miller's. They accuse the military of inertia and of being too ready to diagnose veterans with psychological problems such as post-traumatic stress disorder. The military responded slowly to reports of health problems caused by the use of the defoliant Agent Orange during the Vietnam War (see *Nature* 364, 373; 1993). It now reluctantly supports research on chemical sensitivity.

Nevertheless military officials claim that they can attribute only 30 of the 6,000 reported cases of Gulf-War-related illnesses to chemical exposure. If chemicals were found to cause widespread illness among troops the implications would be enormous. "Battlefields are full of chemicals", says Colonel Rick Erdtmann of the Army Surgeon General's Office. "I'm not sure how the military, or any industry, would function without them."

Susan Greene

'Gulf War syndrome'

Since returning from the Gulf thousands of veterans have complained of symptoms including fever, headache, loss of short-term memory, deteriorating vision, shortage of breath, coughing, diarrhoea, skin changes, bleeding gums, loss of hair and teeth, numbness, tingling, aching joints and fatigue.

In the absence of any diagnosis, the condition has been dubbed 'Gulf War syndrome': a blanket term for one or several illnesses.

S.G.

MITI's research arm changes tack

Tokyo. Japan's Agency of Industrial Science and Technology (AIST), the main research arm of the powerful Ministry of International Trade and Industry (MITI), has introduced a new policy that gives greater freedom to its 15 institutes in selecting individual research projects.



AIST's Tsukuba institutes: major centres favour MITI's change of emphasis.

The move is unprecedented for Japan, where a web of government bureaucracy runs throughout the government research system, and government officials cling jealously to their decision-making powers.

In general, AIST's larger institutes have welcomed the greater autonomy they are being offered. But some of the smaller ones

are hesitating to accept the greater responsibility involved, fearing that they could lose out to more powerful colleagues.

The new policy was agreed last September, and described to the heads of the agency's institutes on 16 July by Hiroshi Kashiwagi, the recently appointed director-general of AIST. (Kashiwagi was previously head of the agency's Electrotechnical Laboratory in Tsukuba.)

From the beginning of the current fiscal year, which started in April, special research funds known as *tokubetsu kenkyuhi* are being set aside by AIST for the institutes to spend as they choose, rather than having to seek prior approval by the agency.

So far, the amount of money involved is not very large, amounting to about ¥2.8 billion (\$27 million) for all 15 institutes. Once personnel costs have been excluded, this represents on average about 10 per cent of their research funds. But

Kashiwagi hopes eventually to double this figure to 20 per cent, a move that would both give greater autonomy to the institutes, and introduce an element of competition between them.

Under the new system, all the institutes put forward proposals for the use of the special funds. The proposals are subse-

quently evaluated by a committee made up of the heads of the 15 institutes' planning sections, which decides how the funds will be allocated.

The chairman of this committee is the head of the research administration division of AIST. But whereas in the past all decisions about which proposals received funding were taken by AIST officials, the chairman's role in the new regime now is simply

Trying to be more international

Tokyo. The new policy to revitalize the institutes of the Agency of Industrial Science and Technology (AIST) also demands greater internationalization. Calls to open up government research are nothing new, but the agency has for the first time committed itself to an ambitious increase in the proportion of foreign researchers. The proposed increase, from 5 to 20 per cent, would add some 300 foreign researchers to the roughly 2,000 now working in the agency's 15 institutes.

The aim is not only to create more fellowships for foreign scientists but also to provide more permanent positions, says Tadatsuna Koda, director-general of the general coordination division of AIST. Most foreign researchers in AIST laboratories are employed on one to two year postdoctoral fellowships. AIST will find it difficult to achieve this, however: visa regulations are restrictive, international schools for children are few, spouses have poor prospects of finding jobs, and language barriers are formidable.

AIST also wants its institutes to publish more in international science journals. Even its best institutes perform poorly compared with equivalents in other countries. For example, the government this week chose as a 'centre of excellence' (see *Nature* **364**, 471; 1993) the National Institute of Bioscience and Human Technology which was formed earlier this year from the Fermentation Research Institute (FRI) and two other AIST institutes. But FRI's 70 researchers published only about 60 papers in international journals in 1992, fewer than one paper per researcher. This compares with around two per researcher per year at Korea's leading research universities, the Pohang Institute of Science and Technology and the Korea Institute of Science and Technology (see *Nature* **364**, 379; 1993).

D. S.

Japan's low-profile 'centres of excellence'

Tokyo. Japan's Science and Technology Agency (STA) unveiled on Monday three 'centres of excellence', to which it will give roughly an extra \$20 million over the next five years. The purpose is to raise the status of the institutions to that of Britain's Medical Research Council Laboratory of Molecular Biology at Cambridge or France's Institut Pasteur. But the world's scientists could be forgiven for never having heard of the three institutes selected.

The three lucky winners are:

■ STA's National Institute for Research in Inorganic Materials in Tsukuba science city, which will study new materials made at 1–10 million atmospheres pressure in the world's most powerful diamond press;

■ The newly established National Institute of Bioscience and Human Technology of the Ministry of Trade and Industry, also in Tsukuba, which will carry out a broad range of research on signal transduction in humans (see *Nature* **364**, 471; 1993);

■ The National Cardiovascular Center Research Institute in Osaka, which will investigate the molecular mechanisms controlling the human circulatory system.

About 20 government institutes from science-related ministries and agencies vied for the first of the new awards introduced this fiscal year. Few details have been given of the procedure used to select the three winners. According to an STA official, each ministry and agency put forward several candidates that were then screened by the "policy committee" (*seisaku iinkai*) of the Council of Science and Technology. The council, nominally Japan's principal science policy-making body, is chaired by the prime minister and draws on the advice of academics, industrialists and civil servants in its various committees.

One criterion was how many papers an institute had published in international journals. STA also considers it important that the centres of excellence should attract scientists worldwide. But none of the institutes chosen has an outstanding publication record, nor are they as renowned outside Japan as some of the other institutes that were not chosen; no foreign scientists were consulted in the selection process.

David Swinbanks

to coordinate the committee's deliberations.

Some institutes welcome the new policy. A leading researcher at the new National Institute of Bioscience and Human Technology in Tsukuba, for example, says that the institute's scientists are "generally happy" with the new system, under which about ¥152 million (\$1.5 million) — 11 per cent of the institute's funds — are now allocated by its director.

But some of the smaller regional institutes outside Tsukuba are less enthusiastic. So far, only a few per cent of their research funds are being allocated under the new system. But these institutes fear that as the system is expanded, they will lose out to their more powerful colleagues in Tsukuba.

In return for the new policy, AIST is asking the institutes to be more open to outside scrutiny in assessing the quality of their research. But there is a general reluctance among the agency's institutes to move in this direction. For more than a year, AIST

has been talking to them about setting up a system of external review, perhaps involving non-Japanese researchers, much as has already been done this year at the Institute of Physical and Chemical Research (RIKEN) and Tokyo University (see *Nature* 363, 570; 1993).

AIST researchers are, however, resisting such moves. Many claim that external reviewers cannot make good judgements about topics in which they are not experts, and that there will inevitably be disagreement between experts in the same field, with the implication that one reviewer's assessment and recommendations could be wrong.

AIST researchers insist that there is no link between the agency's policy of offering greater autonomy to the institutes in funding decisions and its proposal that they should accept external review. But it seems to be AIST's intention that in the long run the two should go hand in hand.

David Swinbanks

Environment institute on the agenda

Washington. The US House of Representatives will consider a bill in the autumn that would create a National Institute for the Environment (NIE) to coordinate better the \$3.1 billion the government spends annually on 'green' research.

Plenty of obstacles litter the bill's path, however: "There are two real questions: turf and money," says David Blockstein of the the Committee for the National Institute for the Environment (CNIE) — a lobby group claiming the support of some 6,000 scientists and other individuals.

But a broad alliance supports the bill. Environmentalists believe NIE would confirm the status of their field. Others see NIE as an opportunity to scrutinize properly the enormous spending on environmental research. "This is about getting extramural research done competently, and properly peer-reviewed," says one Congressional aide. "It's about getting the process out of government."

"Conservative members of Congress recognize environmental decisions will be made, and made either on the basis of political whim or sound science", says Peter Saundry, CNIE's acting executive director. They view the NIE, he says, as "way more science-driven, and less politically-driven" than existing channels.

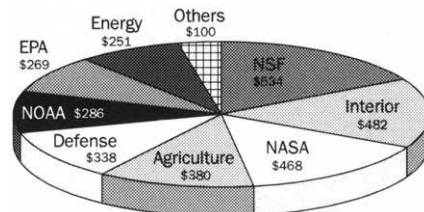
The bill calls for an independent agency free from departmental control. As such it would resemble the National Science Foundation (NSF) rather than the National Institutes of Health (NIH) which are part of the health department. But unlike NSF, NIE's governing board would include not only scientists but also representatives from industry and environmental groups — "multi-stakeholders" as Saundry calls them. The president would select board members on a

rolling basis.

NIE would do extramural research, assess issues for government, provide information and sponsor education and training. Although it would take an interest in all environmental research, NIE supporters say it would not try to grab all of the \$3 billion. Instead it would fill the yawning gap between research and policy. This would en-

US federal spending on environmental research

Total spending \$3,108 million. Individual figures in US\$ million.



EPA: Environmental Protection Agency; NASA: National Aeronautics and Space Administration; NOAA: National Oceanic and Atmospheric Administration; NSF: National Science Foundation
Source: Federal Funding for Environmental R & D, AAAS, 1992

sure that the US body politic got the answers it needs to make informed decisions on everything from laws on vehicle emissions to nuclear dumping.

As it stands, the bill does not mention money. CNIE is reluctant to speculate about what budget NIE would need. "We don't want to raise the red flag of finance too early", says Saundry. He reckons though that just \$30 million could get NIE off the ground, doing "useful assessment work" for government. To start an extramural research programme would need something closer to \$100 million. It would also involve transferring resources from existing agencies.

The difficulties in funding a new agency

were one reason the National Academy of Sciences did not recommend NIE as the best way forward in a report it released in June. "We don't oppose the NIE proposal," says Al Lazen of the academy staff. "We just don't think it goes far enough to solve the problem."

The proposal to create NIE is only one part of a complex debate over about what should be done about the environmental research programme. This is highly fragmented (see pie-chart) and its results satisfy neither side of the increasingly-heated environmental debate. But opinions differ as to whether a new co-ordinating agency is necessary, and if it is, whether it should sit independently or in a government department.

Organization of environmental research is a mess, despite the amounts of money spent, the academy report acknowledges. It recommends at the very least that existing agencies undergo 'cultural change'. It would prefer that a new research-led Department of the Environment incorporate the work of several agencies. A bill going through Congress would elevate the Environmental Protection Agency (EPA) to a department of state.

The academy report failed to bite the bullet, say some NIE supporters. "It didn't go to the heart of the issue," says one congressional aide, repeating a common lament: "They didn't want to trample on people's toes." Saundry says the much-criticized EPA — founded by former President Richard Nixon in 1970 — testifies to the risk of cobbling together new bodies from old ones: the parts, he says, never coalesced into a recognizable whole.

The administration has not commented on the question of how to reorganize environmental research. But its predecessors have created brand-new agencies only where the political imperative was clear-cut: for example, the perceived success of government-funded science during the Second World War propelled NIH and NSF into life, while NASA was created in 1958 in reaction to the Sputnik programme. On the face of it the chances of creating a heavyweight government agency for the environment do not look good. "Right now, economic issues dominate politics", concedes Saundry. "That will pass. This is a good proposal, but it may take time."

The bill is, however, co-sponsored by George Brown (Democrat, California), chairman of the House of Representatives' Science, Space and Technology committee, James Saxton (Republican, New Jersey) and 40 other House members. Brown's support assures the bill hearings, and means it has a realistic chance of being passed in the House — unlike most of the thousands introduced to Congress every year. That would be the first step to establish a NIE: support in the Senate, and funding from somewhere, would still be required to make the project fly.

Colin Macilwain

Irish government turns its back on science

Dublin. After announcing that it would not be funding any new research projects this year, the Republic of Ireland's major grant agency, Eolas, has now found a small amount of cash for its basic research programme. The money will be made available next month. But scientists say it is not sufficient to meet the needs of the research community, and complain that government support for university research remains at rock bottom.

As a result, hopes for building a strong basic research community continue to dwindle in a country with no real science policy and, at less than one per cent of gross national product, one of Europe's lowest levels of spending on research and development. Ireland has no ministry with specific responsibility for science, and major policy initiatives come from industry.

Apart from biomedical research, which is funded by the Ministry of Health, all research is funded by the Ministry for Industry and Commerce. The ministries' funds are distributed through Eolas's basic and strategic research programmes. Researchers are well aware that any grant application should be able to demonstrate a downstream commercial interest; in the case of Eolas, this requires confirmation by a relevant company.

Biomedical research lost nearly half of its funds in 1989 when, government cuts meant that its budget was reduced from IR£2.3 million to IR£1.3 million. Since then, spending on research has been climbing back slowly, and is due to reach £1.78 million this year. But much more is needed, says Vivian O'Gorman, executive director of the Health Research Board.

Funding for Eolas has also been falling, dropping from IR£2.16 million in 1991 to IR£1.85 million in 1993. The latter figure was only just enough to honour continuing commitments, says Martin Lyes, manager of research and technology support at Eolas — hence the agency's announcement in the spring that it was not able to support any new research projects this year. Many scientists were angry that they had wasted time in replying to a call for applications put out by the agency, which closed in January. They have been even more upset that research is being left to stagnate.

John Corish, dean of science at Ireland's oldest university, Trinity College Dublin, says he regrets in particular the freeze on grants for strategic research. "It is a serious blow to the efforts of the universities to bring their research into the market place", he says.

Earlier this month, Eolas announced that an extra IR£150,000 had been found for some basic research projects and will be distributed in September. The money has come from a combination of savings within the agency and contract income. Lyes says the agency has done what it can to help

researchers, but that its options are restricted by its limited budget.

Frustration encouraged researchers to join together to form the Irish Research Scientists Association (IRSA) in May this year. The association, which has about 200 members from all over Ireland, plans to lobby against what it describes as the government's 'antiscience' attitude. Particularly painful, its members claim, is the dismissive attitude of the Minister for Commerce and Technology, Seamus Brennan. In June, Brennan was quoted in the *Irish Times* as saying that the IR£23.4 million in "structural funds" received by Ireland from the European Communities would be restricted to companies, not to "some pet project in a university".

"The whole need for science has been questioned", complains Michael Hopkins, senior lecturer in physics at Dublin City University and a member of IRSA. He welcomes the reintroduction of new research money from Eolas, but says it is too little;

"IR£150,000 is completely inadequate to fund the physics, chemistry and biology projects in Ireland". And it also comes too late to be spent this year.

Frustration among scientists is not restricted to low budgets. Researchers also complain that they seldom know what is happening in government circles, either about policy or available money. Requests from IRSA for briefings with both the minister and with Eolas have not been successful.

Yet big changes are in the air. Eolas is being dismantled in October in a merger between industrial agencies. It will be replaced by a body called 'Forfás' which will have two agencies: the Industrial Development Agency, and Forbairt, which will administer the grants currently handled by Eolas.

In Irish, *eolas* means knowledge and *forbairt* means development. For researchers, the change in agency names may be seen as a symbol of the reality of their situation.

Alison Abbott

Extra funds for Northern Ireland

Belfast. Northern Ireland's two universities have received an extra £4.7 million (\$7.0 million) from the British government for the next financial year to supplement the funds they would otherwise have received purely on the basis of their research performance.

Queen's University in Belfast, and the University of Ulster — which is spread over four main campuses — needed the top-up to compensate for their relatively poor performance in the United Kingdom's recent research assessment exercise.

Over the past two years, the UK government has introduced new criteria for funding research through the university funding councils. These now place more emphasis on research performance, and are intended to eliminate the previous practice of basing research funds primarily on the number of undergraduates (see *Nature* 362, 4; 1993).

Last year, Queen's University received a top-up of nearly £600,000 and the University of Ulster nearly £1 million. Without this, the two universities would have seen their funding for research cut by 0.3 per cent and 3.5 per cent, respectively, while total research funds for higher education in the rest of the United Kingdom rose by more than 11 per cent.

This year, the failure of the universities to improve significantly on their overall research performance would have meant that they were hit even harder. But additional money made available through 'Development Research' funds has avoided the need for more cuts.

This money has been allocated by the Northern Ireland Higher Education Coun-

cil, on the basis partly of research assessment ratings, but also on whether research funds would improve the economic, social and cultural life in the province, and would make the universities able to qualify for funding from other sources, such as the European Commission, whose grants must be matched with a university's own funds.

The two universities had to apply for the money by presenting broad plans for improving their research infrastructure. Queen's subsequently received an extra £2.93 million and the University of Ulster £1.75 million. Each intends to "back winners" by giving most of the money to departments that did well in the research assessment exercise. At Queen's, for example, the physics and electronic engineering departments received top ratings, and will be rewarded with extra funding from the university.

Goodwill is often the victim of hard times. Both universities have benefited from the development research money. But Queen's University also complains that it is being required to subsidize the University of Ulster, which has been faring particularly poorly.

The latter received additional funding from Ulster's limited pool of resources on the basis that, without a 'safety net', its total income would have fallen by more than one per cent, a situation the government has promised to avoid. The higher education council hopes to ease some of the tension between the two universities by giving priority to research projects that promote links between them.

Alison Abbott

Italian health sector in disarray following more scandals

Munich. The Italian pharmaceutical industry is on the brink of collapse following revelations last month of massive corruption in the registration and pricing of drugs. Former health minister Francesco De Lorenzo is under at least ten different mandates for arrest, and the scandal embraces the entire pharmaceutical industry and several university departments.

Giovane Marone, De Lorenzo's secretary, alleged last month that Lorenzo and his colleagues at the ministry — 'the Group of the Golden Pill' — arranged for pharmaceutical companies to pay bribes, or 'contributions', to guarantee registration of their drugs and 'suitable' market prices. The corruption network, he claimed, embraced every link in the chain between the manufacture and marketing of drugs. Its extent and the cynicism of its workings have shocked a nation hardened to tales of corruption.

Since the revelations, police have arrested Duilio Poggiolini, chairperson of the national commission for the registration of drugs and president of the European Commission's Committee for Proprietary Medicinal Products, and Antonio Brenna, chairperson of the national commission that sets drug prices. The government has suspended both commissions.

Elio Rondanelli, president of the Italian AIDS commission, was among the university professors arrested. Antonio Vittoria, who worked at the faculty of pharmacy in Naples, committed suicide after the first allegations. Academics were lured into the corruption web as 'independent experts' to the government and drug industry. They received a share of the 'contributions' when their advice supported the ministry.

Of those charged, only De Lorenzo is free, because he has political immunity as a member of parliament; but he will lose this in the forthcoming general election.

Italy's pharmaceutical companies are accused of collusion, and seeking to influence registration and pricing of their products. One of the most prominent industrialists implicated is Francesco della Valle, who was fired as director of Fidia last year following financial difficulties (*Nature* 364, 562; 1993). Fidia's major product, gangliosides, mysteriously found its way back onto the Italian market only a few days after it was withdrawn in Italy and several other European countries because of its association with Guillain-Barré syndrome, a type of paralysis.

Some companies claim that they were simply victims, not perpetrators, of corruption. Alberto Zambon from the Zambon Group, claims that when his company asked Marone to increase the price of its product Fluimucil in October 1990, he asked it for

IL2 billion (around US\$1.2 million), far more than it was prepared to pay. When the company consequently ran into financial trouble a year later, it was saved from bankruptcy when the Department of Industry stepped in and authorized a small price rise.

Italy's national formulary has long taken the country for a ride. Many of the drugs are useless, and exorbitantly-priced drugs — mainly reimbursable by the state — have put increasing strain on the health budget. Gianni Tognoni, a clinical pharmacologist at the Mario Negri Institute in Milan, estimates that the state could save \$3 billion simply by stopping the sale of just five popular 'but useless' drugs.

Tognoni has also fought against the flaws in Italy's drug policy but without success. It was asking for corruption, he says, because its procedures were so lax. Registration procedures were sufficiently ambiguous, he claims, that drugs could be licensed without documented proof of their efficacy as demanded by law.

Drugs registered for one indication could also be prescribed with relative ease for other, inappropriate, indications. The same drugs would also often be priced differently depending on the company. For example, Guidotti's antihypertensive Captopril cost 30 per cent more than the version sold by the US company Squibb. "At the very least this indicates some confusion on the part of the 'expert' drug price fixing committee", says Silvio Garattini, director of the Mario Negri Institute. As is now clear, high prices were for sale.

Mariapia Garavaglia, the new minister for health, is setting up a new drug registration commission. She will chair the commission, which will comprise 12 independent experts (seven elected from the regions and five by the ministry), the director of the ministry's pharmaceutical department and the director of the Higher Institute of Health — the ministry's technical body. One of the commission's first tasks will be to review the national formulary list and eliminate products with no proven efficacy. It will also introduce new methods of reimbursement to reduce the cost to the ministry of drugs not considered to be essential.

Garavaglia will not set up a new drug pricing commission; the system is too open to corruption she says. She prefers a system of free pricing where industry decides on the cost of its own products. The Department of Industry opposes such a move, and has set up a group of experts to advise it on the criteria on which drugs should be priced. It believes the government should retain control over drug costs as it reimburses much of the country's drug bill.

Alison Abbott

Brazil's proposed law threatens animal experimentation

São Paulo. Scientists in Brazil are up in arms about a proposed animal experimentation law which they fear could threaten the future of biomedical research there.

Critics of the proposed law, which was prepared for the government by the Brazilian lawyers' association, concede that legislation is needed to end the existing free-for-all in animal experimentation in the country. But they object to the proposed structure, under which regulation will be controlled by a powerful, autonomous committee whose membership rules will explicitly exclude any scientist who has recently been involved in animal experiments.

"It has elements of extreme madness," complains Mauricio Rocha e Silva, head of the Divisão de Experimentação (Experimentation Division) at the prestigious Incor (Instituto do Coração, or Heart Institute, affiliated with the Faculty of Medicine of the University of São Paulo).

Rocha e Silva chairs a committee recently created by the Brazilian Academy of Sciences and the Federation of Experimental Biology Societies to voice scientists' views on the subject. "The project conflicts with scientific freedom," says Rocha e Silva, besides "creating a sort of a police that will stop research by force."

Circe Amado, author of the draft of the new law, does not agree. An environmentalist lawyer, Amado is engaged in a project to consolidate Brazilian laws dealing with the environment. The project deals not only with use of animals for scientific purposes, but all other instances of possible cruelty to them.

Amado denies that the project will hinder research. She says that scientists have become insensitive to animal suffering and defends the right of inspectors to close down laboratories whenever any kind of animal cruelty is detected. According to Amado, the project is based on equivalent legislation in both Britain and Sweden.

The focus of scientists' anger is a proposed organ called COSETICA, or the Comitê Superior de Ética Científica com Animais (Superior Committee for Scientific Ethics with Animals), which would have the power to close down labs in case of perceived cruelty, and to license researchers to use animals.

The project also suggests that COSETICA should have "autonomy," freeing it from the influence of the science ministry or other government departments. "It is a republic inside the republic," complains Rocha e Silva; "it is against the Constitution." Amado answers that up to now Brazil has had no animal rights law and that one with enough power is needed.

Ricardo Bonalume

SDC truly international

SIR — As spokesman of the Solenoidal Detector Collaboration (SDC), I am responding to the news item (*Nature* 362, 385; 1993) about non-US funding for the SDC detector project. The SDC is an international collaboration involving more than 900 physicists and engineers, many of them from overseas, that has been working intensively over the past three years on the development of a detector intended to exploit the physics opportunities opened up by the Superconducting Super Collider (SSC), now under construction near Dallas.

There are several errors in your article: Professor "Ott" from Canada is really Professor Robert Orr, the National Science and Engineering Research Council is really the Natural Sciences etcetera, and the statement attributed to the French science attaché that "French scientists now working on the SDC are being paid by the SSC" is incorrect.

But I object most strongly to the negative tone of the article, well reflected in its title, "SSC detector collaborators shun financial commitments". It is correct that the non-US funding for the SDC detector, expected to be about 40 per cent of the total, is not yet firm. However, this is not surprising. Before specific requests can be made to funding agencies, the detector design and research and development must be sufficiently far advanced that one can define with some precision the in-kind contributions to be made by the non-US collaborating groups. The SDC is just reaching that stage, after successfully completing two intensive reviews, including one by a distinguished international group of scientists. The specific dollar numbers quoted in your article represent the value of the items that our non-US collaborators are hoping to provide, estimated according to US accounting rules. The SDC is developing specific country-by-country proposals which will, at the appropriate time, be submitted to the relevant funding agencies.

The tradition of international collaboration in the construction and exploitation of detectors in high-energy physics, with involvement much broader than just the country or countries financing the accelerator operation, is well established. There are unique facilities in Europe (LEP, HERA) and in Japan (TRISTAN), and American groups have been actively involved in experiments at these facilities with financial support from the US government, and in collaboration with groups from Europe, Japan, Canada and elsewhere. Similarly, some US facilities such as the Tevatron Collider at Fermilab provide unique capabilities, and groups from outside the United States are collaborating with US groups on experiments there.

The 40-TeV collision energy of the SSC will unquestionably open up unique scientific opportunities not available anywhere else in the foreseeable future. Participation in the SSC experimental programme by interested groups from outside the United States is therefore the natural continuation of a well-established and highly successful tradition.

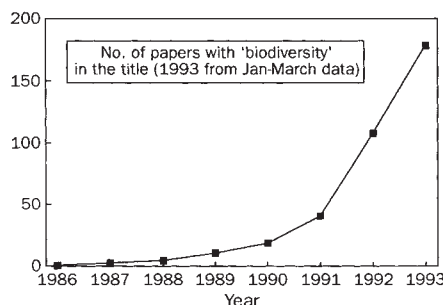
George H. Trilling

*Lawrence Berkeley Laboratory,
University of California,
Berkeley, California 94720, USA*

■ Professor Trilling is correct in saying that French scientists working on the project were not paid salaries by SDC; we were misinformed by the French Embassy in Washington. — Editor, *Nature*. □

Overbiodiverse?

SIR — While recently carrying out a literature search on the BIDS (Science Citation Index) database I was struck by the incidence of papers whose titles contained the word 'biodiversity' (see figure). Since the first record in 1987, and estimating the current year's total from the first three months, it is apparent that the rate of increase of biodiversity publications (by this criterion) is almost exactly exponen-



tial, with a doubling time very close to one year.

Does this trend indicate a growing awareness of the problem of defining and conserving biodiversity, or a tendency for the discussion of problems to substitute for the ability to solve them?

Peter J. Hogarth

*Tropical Marine Research Unit
and
Department of Biology,
University of York,
York YO1 5DD, UK*

Epidemic misuse

SIR — We are distressed at the tendency to use terms referring to human diseases to describe those of nonhuman animals. Thus we have "epidemic" (for 'epizootic')¹⁻⁶, "epidemiology" (for

'epizootology')^{6,7}, "symptom" (for 'sign')⁴⁻⁹ and "autopsy" (for 'necropsy')¹⁰.

The derivation of the terms makes it plain that these are misusages. The particle *-dem-* derives from the Greek *demōs*, meaning 'people' (whence also 'democracy'). A 'symptom' requires communication with a human patient, but a 'sign' of disease is evident by observation. Although the derivation of 'autopsy' does not restrict its use to people, the standard definition¹¹ does.

The mixing of these terms in the same publication succinctly illustrates the confusion. Why are mass mortalities of sea urchins² in California and Canada "epizootics", but those in the Caribbean an "epidemic"²? Is there⁶ an 'epidemiology' of fish epizootics? Signs of 'ecosystem distress' there may be⁶, but can there be 'signs and symptoms' thereof? What¹² is the 'epidemiology' of 'epizootic apthae'?

We do not suppose that these mistaken substitutions are deliberate. They may be consequences of ignorance. Or are they manifestations of political correctness modulated by the view that there is no proper distinction between human and nonhuman animals?

Ernest H. Williams Jr, Lucy Bunkley-Williams (*Marine Ecological Disturbance Information Center, Department of Marine Sciences, University of Puerto Rico, PO Box 908, Lajas, Puerto Rico 00667-0980, USA*); **John M. Grizzle** (*Department of Fisheries and Allied Aquaculture, Auburn University, Alabama 36689, USA*); **Esther C. Peters** (*Tetra Tech Inc. 10306 Eaton Place, Suite 340, Fairfax, Virginia 22030, USA*); **Donald V. Lightner** (*Department of Veterinary Science, University of Arizona, Tucson, Arizona 85721, USA*); **John Harshbarger** (*Registry of Tumors of Lower Animals, NHB-163, Smithsonian Institution, Washington, DC 20560, USA*); **Aaron Rosenfield** (*Northeast Fisheries Center, Oxford Laboratory, Railroad Avenue, Oxford, Maryland 21654, USA*); **Renate Reimtsma** (*Department of Pathology, Medical School Teaching Facility, University of Maryland School of Medicine, 10 South Pine Street, Baltimore, Maryland 21201, USA*)

- Gould, S. J. *Nat. Hist.* **6/91**, 22 (1991).
- Lessios, H. A. *et al. Science* **226**, 335 (1984).
- Rowan, R. & Powers, D. A. *Science* **251**, 1348-1351 (1991).
- Lessios, H. A. *et al. Coral Reefs* **3**, 173 (1984).
- Lessios, H. A. *et al. Rev. Ecol. Syst.* **19**, 371 (1988).
- Cairns, V. W. *et al.* (eds) *Contaminant Effects on Fisheries* (Wiley, New York, 1984).
- Peters, E. C. *Helgol. Wiss. Meeresunters* **37**, 113-137 (1984).
- Grigg, R. W. *Coral Reefs* **11**, 183-186 (1992).
- Muehlstein L. K. *Can. J. Bot.* **70**, 2081-2088 (1992).
- Novotny, J. F. & Beeman, J. W. *Progr. Fish-Cult.* **52**, 162-170 (1990).
- Post, G. & Kiontz, W. G. (eds), *Glossary of Fish Health Terms* (Amer. Fish. Soc. Fish. Health Section, 1977).
- Siegmund, O. H. *et al.*, *The Merck Manual* (Merck & Co., Inc., Rahway, New Jersey, 1979).
- Anderson, C. *Nature* **351**, 89 (1991).
- Nature* **345**, 648 (1991).
- Webster, R. G. & Yuanji, G. *Nature* **351**, 527 (1991).
- Weissmann, C. *Nature* **349**, 569-571 (1991).

Competition in science

SIR — Most thoughtful observers would agree that “excessive competitiveness” can have many harmful effects on both scientists and science itself (*Nature* 363, 667; 1993). However, excessive competition in science may claim a casualty in addition to those specifically discussed in your article, that is, truth itself. The rush to claim priority may result in the truncation of efforts to establish the validity of the observations that created the excitement in the first place. Yet no one wants the embarrassment of publishing an “exciting” paper which is later shown to be incorrect.

Stephen J. Gall

*Departments of Pathology,
Beth Israel Hospital
and Harvard Medical School,
Boston, Massachusetts 02215, USA*

SIR — John Maddox worries, as others have done before, that competition in science is becoming too fierce. Those of us who have endured recently the research selectivity exercise in the United Kingdom will react with hollow laughter to his suggested remedy, sensible though it is, that the link between publication record and reputation should somehow be broken.

Maddox gives details of one aspect of competitiveness, the rush to get into print, which has all the decorum of the opening of the annual sale at a department store. He mentions also the connection with fraud in science. But there are other disadvantages to too much competition. There are concerns about the unit of currency in science; Rosner¹ drew attention to the increasing prevalence of “assertive sentence titles”, which reduce the content of a paper to its title. Crichton² worried that the writers of review articles are not thorough enough, and in doing so alluded to the most serious effect of over-competitiveness, the erosion of time to think.

The scientist was once an all-round thinker, but is now reduced to a searcher for the next item of information. The amount of published work increases rapidly and it is sometimes impossible to keep up in one's own subfield, let alone read more widely. The context or consequences of one's work become subordinate to the facts. There will always be some who prefer to work with this approach, but the present circumstances are forcing it to be the only approach. With that approach, this letter and all comments on the process of science become subordinate to the science itself. I feel this cannot be for the best; yet if asked for evidence it is difficult to set up a testable hypothesis with which to provide it. It is partly this apparently blind rever-

ence for facts that has prompted the recent books attacking science^{3,4}.

The number of biomedical journals is increasing exponentially⁵, but much published work is never cited⁶ and Hamilton quotes a professor at Massachusetts Institute of Technology saying: “If the bottom 80% of the literature just vanished, I doubt if the scientific enterprise would suffer.” There is immense waste in science. How much waste is necessary to ensure that real progress is made? Does science really have to proceed at the pace it is being pushed? And how much of the waste is *because of* competition?

Neville W. Goodman

*University Department of Anaesthesia,
Southmead Hospital,
Bristol BS10 5NB, UK*

1. Rosner, J. L. *Nature* **345**, 108 (1990).
2. Crichton, R. R. *Nature* **337**, 110 (1989).
3. Sutherland, S. *Nature* **357**, 550–551 (1992).
4. *Nature* **358**, 698 (1992).
5. Wyatt, J. *Lancet* **338**, 1368–1373 (1991).
6. Hamilton, D. P. *Science* **250**, 1331–1332 (1990).

Insurance risks

SIR — I am perfectly comfortable with the arguments advanced about the ethical issues involved in genetics with one exception. You ask “why should [insurance companies] also not discriminate against, say, people with the particular structure of the LDL receptor known to be responsible for early-onset heart disease?”¹ In fact, however, there is an enormous literature available in the legal world on why it is unwise to allow insurance companies to set their rates on the basis of immutable characteristics such as race or other unchanging attributes (see, for example, refs 2–4); indeed, the various laws of the United States frequently prohibit insurance companies from using this type of data when setting rates.

Essentially it is argued that the essence of insurance is to spread risks of certain illnesses and diseases throughout society and not to burden those who bear the exclusive risk with its full costs. This social policy should seem to prevent precisely that type of rate setting data from being used. This type of law is quite distinguishable from the legally permissible policy of raising rates for those who engage in voluntary pleasurable activity that increases their risk of illness, such as smoking.

Michael Brody

*Department of Religion,
Emory University,
Atlanta, Georgia 30322, USA*

1. *Nature* **364**, 97 (1993).
2. Lowe, R. *Drake Law Rev.* **40**, 507 (1991).
3. Miller, J. *Dickinson Law Rev.* **93**, 729 (1989).
4. Kolata, G. *Science* **232**, 317 (1986).

Cite unsound

SIR — Like most other scientists, from time to time I receive papers to review. It has been my practice to check all the references that I can easily get hold of, checking that both the bibliographical information and also the cited data or opinions are correct. If a paper has been revised several times, it is almost inevitable that the interpretation or context of some citations will change (and should be checked before anyone else sees the paper). However, I have been surprised at how many citations are just plain wrong.

For example, a citation to insect damage never mentioned insects in general or in particular. A paper cited as comparing grazing versus hay production consisted entirely of hand-cut plants grown in pots in a controlled environment chamber, while one that was claimed to show the effects of phosphorus and potassium fertilizer only looked at nitrogen. I have even had someone quoting himself as finding the opposite from what he actually found. A colleague tells me he was once asked why he did not quote the standard reference on a topic. His response was that, despite being widely quoted, it did not exist.

Is careless citation common to other areas of research or is it peculiar to agronomy? How do authors expect us to trust their data, which we cannot check, if their citations, which we can check, are treated so cavalierly? I urge all authors to check references carefully, and reviewers to reject papers with careless citation, to try to cut down on the amount of misinformation entering the system.

R. R. Symons

*Agriculture Canada,
Brandon Research Station,
PO Box 1000A, RR3,
Brandon, Manitoba,
Canada R7A 5Y3*

Fetal tissue

SIR — The Commentary article by Bianchi, Bernfield and Nathan (*Nature* 363, 12; 1993) is mistaken regarding the scope of the five-year moratorium on US federal support for fetal tissue transplantation research. The moratorium curtailed the use of electively aborted human fetal tissue in therapeutic transplantation research, but it did not stop the use of human fetal tissue in basic research. In fact, during those five years, the National Institutes of Health funded approximately \$8 million annually in basic research using human fetal tissue. Human development and cell differentiation were topics of a number of those studies.

Sarah Carr

*Science Policy Studies Center,
National Institutes of Health,
Bethesda, Maryland 20892, USA*

Making more of academic assets

Nicholas Scott-Ram

Why is the UK's biotechnology industry still nascent? Academic institutions must exploit the commercial opportunities of the research they support if Britain is to remain internationally competitive.

ACCORDING to almost every criterion — research funding, income generated (\$8.1 billion versus \$525 million in 1992), the amount of capital invested and the number of companies that have gone public — the UK biotechnology industry lags far behind that of the United States. Even now, only a handful of biotechnology companies are established in Britain, although that number is gradually increasing. What is most disappointing is that the high quality of basic, academic research in universities has not been translated into commercial opportunity.

What makes a successful biotechnology industry? Apart from essentials such as a high-quality research base, access to finance and regulations covering research and manufacturing, product approval and so on, two important factors are efficient technology transfer and patent protection by individuals, academic institutions and industry.

Challenges

Technology transfer covers a range of activities including the transfer of information, know-how, intellectual property, materials, potential products and processes from academia to industry. The gap between industry and academic institutions is often referred to as the 'development gap', and the challenge for effective technology transfer is to find ways of bridging this gap, whether it be through licensing agreements or establishing spin-off companies.

Technology transfer has been particularly successful in the United States in health care and related areas (see table). For example, the osteoporosis agent Calcifedil, invented at the University of Wisconsin, has had worldwide sales of around \$400 million. Another osteoporosis agent, Calcitriol, had worldwide sales of \$35 million. The anticancer agent Cis-platin was discovered at Michigan State University; annual revenues are in the region of \$150 million. A burn ointment, Sulfadiazine, discovered at the University of Columbia, is another \$100-million product, whereas Warfarin, again from the University of Wisconsin, has revenues approaching \$20 million per annum. The original Cohen-Boyer patent of Stanford University covering the basic technique of genetic engineering is expected to earn over \$100 million in licence fees by the time the patent expires.

Why is there a substantial gulf between

the United States and the United Kingdom in the success rate of the commercialization of academic research? Technology transfer in many universities in the United Kingdom is of a highly variable nature. All too often, high-class university research stands little or no chance of being commercialized. Several institutions, such as Oxford, Imperial College and University College London, have an excellent reputation for technology transfer, but there are still many centres of academic excellence which lag behind. If inventions are not commercialized, then everyone loses out — not just the inventors, but also the universities, industry and the country's research base.

Technology transfer is a percentage game (the percentage varies from one discipline to another), where for every success story, there are a series of failures. Consequently, technology transfer, if it is to obtain a satisfactory return, requires adequate financial resourcing and management expertise. Success stories in the United States such as Stanford University, which had a licensing income in 1989 in excess of \$13 million, and MIT (\$6.5 million licensed income), have resulted from the development of sound strategies for technology transfer and the realization that the pay-off takes a very long time: financial returns may not happen until 10 years after patenting.

Central to technology transfer is the handling of intellectual property, including patents, copyrights, trademarks, registered designs and trade secrets. Patenting is probably the most important. First, a granted patent will protect an invention for up to 20 years, which allows the patentee to prevent third parties from practising the invention without rewarding the inventor. Second, in the early stages of technology transfer, whether it be by a licence arrangement or by setting up a new company, one of the key assets to investors are the patents. For example, venture capitalists usually seek investments where there are patents. The London Stock Exchange requires a biotechnology company to have granted patents before allowing it to be fully listed on the Stock Exchange. Recent disputes within the biotechnology industry have starkly demonstrated the value of patents, such as the Amgen v. Genetics Institute case on erythropoietin (EPO), and the Genentech v. Wellcome case on tissue plasminogen activator (tPA). In the case of Amgen,

worldwide sales of EPO are now approaching \$500 million, while the loser, Genetics Institute, was subsequently acquired by American Home Products.

All this means that universities, if they are to be involved in technology transfer, must pay much more attention to patenting and the mechanisms for technology transfer. 'Prosecution' is a process involving responding to the patent examiner's official actions and convincing him or her that the application is worthy of a patent. The filing and prosecution to grant of a patent application on a worldwide basis can cost anything between £30,000 and £100,000 over a 5- to 10-year period. This is a substantial investment and, while it need not all be met by the university, universities should be prepared to provide realistic budgets for patenting. A sensible compromise is for a university to seek general protection by filing a preliminary PCT (Patent Cooperation Treaty) application, which gives protection for several years without incurring substantial costs before national examinations. At the same time, the risks involved in these levels of financial investment can be reduced by careful preparation. Undertaking literature searches, including patents, before filing will at least give an indication of what the competition is doing.

It is particularly important for biotechnological inventions that universities should use patent agents who are expert in this field. There is a substantial difference between inventions in the engineering and biotechnology fields, and requiring a patent agent experienced in the one to draft sound claims for the other is asking too much. In addition, universities should enlist patent agents who have had experience of patent filing and prosecution in the United States, where the rules and tactics differ greatly from the United Kingdom.

Strategies

It is common in academic circles for a patent application to be filed at the last moment before publication, a practice not favoured by commercial organizations. Patent applications filed in a hurry with poor exemplification can jeopardize the future scope of the invention and make prosecution difficult and expensive. Although academic scientists are reluctant to delay publication, more thought about patenting before submission of an article for publication can save later problems. Another area which is sometimes

overlooked by academics is the question of whether grant applications are citable as prior art. In the United States, grant applications are regarded as confidential until awarded, and thereafter they are in the public domain and thus citable as prior art against a patent application. In Europe the situation is similar in that if a grant application becomes public knowledge, then it is prior art. While it is unlikely that an invention will be disclosed in a grant application, often "problems to be solved" are set out, along with the current state of the art, which could be cited against an application. Once the patent application has been filed, universities must be prepared to invest the time and effort required to prosecute their applications. For example, interviews with patent examiners, who often favour academic over commercial patentees, can help the prosecution process.

The need for effective patenting strategies is critical in the light of international competitors. Patenting of biotechnological inventions internationally presents a special challenge in its own right. Difficulties have been encountered in interpreting biotechnology patents within the existing framework of patent law, both in Europe and in the United States, which has led to inequality in the scope of granted claims for many inventions. Even within each country's national patent office there are inconsistencies in examination and the scope of claims granted.

These differences in national patenting requirements are also amplified because of the subject matter involved: biological entities and processes or, more specifically, proteins, genes, vectors and cell lines, where there remains a significant knowledge gap between the science and the law. The assumption is made that such inventions are similar to chemical patents, and that no new rules are required to define their patentability. Decades of chemical patent practice have produced generally accepted conventions concerning the balance to be struck between actual exemplification of an invention and the scope of an allowable claim defining the invention. In addition, structure-activity relationships are often more clearly discernable with the small molecules of traditional medicinal chemistry. Such a corpus of knowledge does not yet exist in biotechnology, and it is likely to be some time before it does. In the United States, the courts have yet to give a definitive ruling on the question of the patentability of a DNA sequence of a previously known protein and the structure and function of novel biological entities is often not defined. A common example is that although natural molecule X was known in the art, a patent covering the generation and composition of molecule X through recombinant DNA technology is often of

sufficiently broad scope to cover all variants of molecule X, even if the properties of the second generation X are substantially different from the natural molecule. One consequence is that patents are filed with broad claims which are unsupported in the specification and are subsequently open to attack. Until recently, this trend was more frequent in the United States than in Europe, although the balance now seems to have reversed.

A further complication is the substantial differences in the patent laws of Europe, the United States and Japan. In the United States, the current priority

commercial organization will want as firm a patent as possible before entering an expensive development programme.

A famous example of the different attitudes to intellectual property between government and academic institutions in the United States and Europe is the US National Institutes of Health (NIH) application for 2,000 novel gene sequences with no known function, otherwise known as the Venter application (see *Nature* 353, 485; 1991). Such genes or fragments have been identified by complementary DNA sequencing rather than genomic sequencing, and the only use claimed in

HOW MUCH MONEY UNIVERSITIES MAKE ON TECHNOLOGY TRANSFER

	1989 R&D expenditures (\$ thousands)	Tech. transfer office budget (\$ thousands)	Licensing income (\$ thousands)
University/research centre			
Stanford University	290	1,583	13,700
MIT	280	1,000	6,500
University of Pennsylvania	160	500	3,000
Cornell University	280	800	2,500
Rutgers University	106	700	1,750
Harvard University	305	1,050	1,600
North Carolina State University	116	130	1,284
Fox Chase Cancer Center	34.2	600	1,200
Tufts University	75.5	95	1,140
Massachusetts General Hospital	100	N/A	850
University of Utah	80	464	632
Yale University	190	200	600
Thomas Jefferson University	25	275	500
University of Maryland, College Park	136	250	250
Case-Western Reserve University	120	330	250
Virginia Polytechnic Institute	90	250	250
Georgetown University	45	266	50
University of Maryland, Baltimore	58	N/A	20

Technology transfer has been particularly successful in the United States in health care and related areas. Source: Innotage Management Inc./Corporate Venturing News, 11 November 1991.

date of a patent is based on a first to invent, rather than first-to-file system. The US system also favours the academic scientist with the one-year's grace rule, which allows publication up to one year before a patent is filed, but in Europe this may invalidate the patent. In Japan, a disclosure in a learned journal within 6 months of the filing date will not destroy novelty. Such grace periods as there are in Europe are drastically restricted in scope by comparison, and this is one area where harmonization between countries is needed. There are also differences over what is patentable subject matter. Even in Europe there are differences in interpretation of patent law between countries. A recent example was the Genentech tPA case where in the United Kingdom the patent was revoked after an action by Wellcome, whereas elsewhere in Europe an opposition hearing upheld the Genentech patent. The same patent has been upheld in Japan. Such uncertainties can make technology transfer difficult as the

the application was that the genes or gene fragments could be used as probes. The application was initially rejected by the US Patent and Trademark Office on various grounds including lack of novelty (as the allegedly novel molecules had been claimed very broadly); but this rejection was only the first stage in the dialogue between the patent office examiner and the application.

The filing of the Venter application highlights several important patenting issues. The early disclosure of sequences with unknown functions may jeopardize the filing of later patent applications which may contain parts of those sequences but which also define a specified function. If a patent is granted for these fragments and they later dominate research, the patent research could potentially slow down the research. There are also concerns that the NIH's Venter filing indicates an increasingly protectionist attitude to biotechnology in the United States. The European Patent Office stated

last year that it would not grant applications for DNA sequences with unknown functions. A reasonable view would be that the patenting of partial sequences of complementary DNA is not a complete invention but really relates to intermediate research results. Until functionality has been defined, there is insufficient data on which to file a patent. Issuing a patent based only on partial sequences would dominate any patents filed later which ascribed a function to these sequences, and this would be harmful to future research.

In Europe, there has also been controversy surrounding the European Community (EC) draft directive on the protection of biological inventions, which attempts to address areas such as the patenting of life forms, particularly animals, and the question of farmers' privileges. The original draft, proposed in 1988, aimed to clarify the law and provide greater certainty in the member states, ensure uniform interpretation throughout the EC and improve the level of protection for both inventors and investors to the levels available in Japan and the United States. This had not happened, and in several major respects, the revised proposals are contrary to provisions in the European Patent Convention, the 1991 International Convention for the Protection of New Varieties of Plants (UPOV), the "Uruguay Round" on GATT (General Agreement on Tariffs and Trade), and the World Intellectual Property Organization's present negotiations towards a patent harmonization treaty.

It is unlikely that the EC draft directive will serve any purpose other than to add another layer of complexity to an already complex situation. If the EC is to remain competitive in biotechnology, then European patent laws will have to track developments in the United States, particularly in areas such as the patentability of life forms. The failure of the draft directive to tackle this problem is significant.

One of the most controversial areas relates to the patenting of life forms. The recent case of the Harvard 'oncomouse' is a case in point. The Harvard oncomouse is a transgenic mouse which has a predisposition to forming tumours and can therefore be used as an animal model for assessing anti-tumour agents, thus speeding the progress of drugs to the clinic. Although the application was originally refused by the European Patent Office, it was subsequently granted after an appeal; the patent is now being opposed by various groups on ethical and moral grounds. The real issue is the competitiveness of the EC with the United States. If the scope of patenting is broader in the United States than in the EC, the EC will be at a disadvantage. Whatever the final position on patenting life forms, there

should be consistency between the principal patenting territories.

Over and above these issues currently relating to the granting of biotechnology patents, there are also problems surrounding the seeding of these patents into industry. When it comes to the search for an industrial partner, many universities fail to invest the time and money required for successful technology transfer. Although an industrial application is a requirement of a successful UK patent application, there remains the difficulty of defining a market and hence a partner. Technology transfer has to be market driven — it is no good patenting an invention and then expecting an industrial partner to come in off the street. The technology has to be sold and this requires an investment of time and effort. Consequently, university administrators with little or no experience in technology transfer are no substitute for experienced professionals.

Encouragement

Once a potential industrial partner has been found, other problems arise. Too often entrepreneurial academics receive little or no encouragement from the university, although the corporate partner will usually need support from the academic in the transfer of know-how. Nor do academics receive much in the way of financial reward if the technology is successfully transferred. Excessive overheads for industrial funding of academic research are common (with figures around 140 per cent of direct costs being quoted), as is a reluctance to address who has responsibility for patenting.

The variety of ways of sharing intellectual property rights and ownership between the university and the industrial partner can present a problem. For example, joint ownership of patent rights between the university and a corporate partner adds to the bureaucracy and is unacceptable to the corporate partner, who generally wishes to be in control of the situation and will not object to paying a reasonable royalty to the university once the product is commercialized.

Given the complexity and urgency of this situation, it is unfortunate that there are issues in the United Kingdom concerning the attitudes of academic scientists to commercialization. Some important barriers to profiting from science in academic institutions appear to be more psychological than legal. Although institutional mechanisms are often in place and available to academic scientists for the commercialization of their inventions and entrepreneurship, British academics are often reluctant, and sometimes hostile, to using such mechanisms. The attitude persists that 'pure' research is superior to 'applied' research, although this distinction is becoming

increasingly outmoded.

Another prevalent factor is a fear of failure. This is odd when it is considered that an academic scientist is in a 'no-lose' situation. In career terms, an academic scientist's success is usually measured in terms of publications, and not on the basis of commercialization of ideas. Within the requirement of obtaining a patent, commercial organizations do not usually discourage scientists from publication.

Part of the problem is the attitude of university administrators. Many believe that industry should pay a premium for being allowed to benefit from university research, but they fail to appreciate the realities of financial investment and risks involved in taking a product to the market. The development of a marketable product from a research project within a university is time consuming, expensive and needs considerable skill — facts which are not at times appreciated by university administrators. Finally, there is a significant school of thought that believes academics should not benefit financially from their inventions on the grounds that it disrupts parity between academics and fosters divisiveness.

How can the situation be improved? First, universities and other academic institutions should focus on protecting the value of their intellectual property rights by educating academic personnel on patenting and commercialization options and through the sensible filing of preliminary patent applications. Although this has been said many times, it needs to be repeated until the message has sunk home.

Second, technology transfer should not be regarded as just another form of university administration. Those responsible for technology transfer in universities should be experienced professionals who understand the needs of industry. Although the initial investment may be higher, the rewards will more than justify it. Cutting corners at the start only increases the chances of failure.

Finally, as regards patenting, the British government should appreciate that moves towards improved harmonization of patents are urgently needed for the United Kingdom to be competitive. Parallel negotiations are underway through the World Intellectual Property Organization and GATT. The slow rate of progress in these negotiations is disturbing and efforts should be made to speed up the process. It does seem that the United States will be moving to a first-to-file system in 1994, but it is likely that concessions will have to be made in Europe in return. The more the UK, EC and US systems are harmonized, the easier it will be for biotechnology to prosper in the United Kingdom. □

Nicholas Scott-Ram is at British Biotechnology Limited, Watlington Road, Cowley, Oxford OX4 5LY, UK.

Putting molecular biology into water

Complaints that molecular biology is indifferent to quantitative considerations are belied by brave efforts to understand the hydration of glucose.

THOSE who complain from time to time (and even often) that modern biology is insufficiently quantitative have a responsibility to look for signs that the wind is changing and then to say so. What follows is in that spirit. It is an evocation of heroic efforts recently made to tackle what may be thought one of the rudimentary problems of biological chemistry, that of the behaviour of the glucose molecule in solution in water. A little reading leaves no doubt of how much effort (and Cray-time) has already been expended on the task. It also confirms that there is much to be learned about glucose before the real polysaccharides with which cells are filled can be understood.

The complainants are fond of saying that one of the glaring deficiencies of molecular biology is that its descriptions of the behaviour of the interesting molecules of biology, from protein enzymes to DNA, almost always overlook their interaction with the sea of water molecules in which they are invariably immersed. The complaint, it goes without saying, is not a moral complaint against the character of those who practise molecular biology. Rather, it is a comment on the inherent difficulties of describing water interactions on a molecular scale. It might be different if there were a calculable theory of water itself. Since there is not, it may seem remarkable that so much can be said about glucose in solution.

Glucose, everybody knows, is a simple molecule with the formula $C_6H_{12}O_6$, which is hardly more complicated than the formula for benzene. Moreover, both molecules have a similar architecture, each being built around a six-membered ring of atoms. The difference is that the six-membered ring of glucose consists of five carbon atoms and an oxygen, and that it is held together by single and not double bonds. To one of the carbons next to the ring oxygen is attached a CH_2OH group — essentially methyl alcohol, while every other carbon atom in the ring carries a hydroxyl group characteristic of alcohols (plus a hydrogen atom to make up the carbon valency of four). The water interactions of the glucose molecule must plainly be central to its real-life properties.

But even in the absence of water, there are complications enough. For one thing, the ring-carbon next to the ring oxygen that carries both a hydrogen atom and a methyl alcohol residue is intrinsically asymmetrical, and so must be a centre of optical activity. (Real-life glucose consists of the D enantiomer.) So is the ring-carbon on the other side of the ring-oxygen (convention-

ally known as C1), which makes four distinct linkages with a ring-carbon, the ring-oxygen, a hydroxyl group and a hydrogen atom.

Worse, being aliphatic, the six-membered ring cannot be geometrically flat, like that of benzene, but must be puckered. In principle, there are three possible directions of puckering, corresponding to the three diameters across the roughly hexagonal ring, for each of which there may be two puckered forms, corresponding to the cases in which the atoms at the opposite ends of the diameters are displaced in the same or opposite directions above (or below) the plane formed by the remaining four atoms (the "boat" and "chair" forms of puckered rings respectively).

It seems common ground among glucosologists that the only form of the glucose molecule whose stability earns it a place in nature's scheme of things is that in which the oxygen atom of the methyl alcohol residue lies roughly in the plane of the puckered ring, which is the D enantiomer observed, and does not project away from it roughly at right angles. The other complication is the configuration of the groups attached to C1, which switch from one form (called α) to the other (called β) naturally in solution. For some time, it seems to have been taken as a mark of progress in the field that there should be an explanation why, at equilibrium, there should be twice as much β as α anomer in solution.

Simple handwaving has done much to keep the field alive. One line of argument has been the observation that intramolecular hydrogen-bonding may stabilize some forms at the expense of others. The idea is that some configurations of the whole molecule can place as many as all six of the hydroxyl hydrogen atoms near to other oxygen atoms, creating for the central ring of glucose a kind of simulation of immersion in water. The same line of argument favours the relative stability of the β -anomer: the C1 hydroxyl group then lies roughly in the equatorial plane of the molecule, and so takes part more easily in internal hydrogen bonding. But that, of course, is only handwaving, and says nothing about external water interactions.

Now Christopher J. Cramer and Donald G. Truhlar from the University of Minnesota, Minneapolis, have given glucose what might be called the full Cray treatment (*Journal of the American Chemical Society* **115**, 5745–5753; 1993). The goal is to deal with all the aspects of the glucose molecule in one

fell swoop — the electronic binding of the atoms together, their relative motion (vibration), the motion of the intact molecules and their interaction with the external aqueous environment. It is a *tour de force*, although no doubt merely a foretaste of what lies ahead.

One interesting conclusion is that the predictions of handwaving are not borne out. What seems to stabilize the β relative to the α -anomer is not so much the intramolecular hydrogen bonding as the vibrational patterns of the two forms. Nobody could have guessed at that by handwaving. But clearly the number-crunching will have to be applied to many of the other isomeric forms of glucose before all the subtleties of that molecule, even in the 'gas phase' (which means no water), are fully understood.

And, for all the care that Cramer and Truhlar have taken, the water interactions remain the weak point of the calculation. Although it would be possible in principle to set up a molecular dynamics routine to deal with even the most ramified chains of interacting water molecules connecting pairs of hydroxyl groups on a glucose molecule, that would be equivalent to solving the whole water problem in especially taxing circumstances. The snag is that it is quite possible that some of the peculiarities of biologically significant molecules in solution derive from the persistence of extended interaction chains through water molecules in the medium.

Instead, Cramer and Truhlar (no doubt for the time being) are forced to rely on the solvation models they have developed in the past few years (see *Journal of Computer-Aided Molecular Design* **6**, 669; 1992), and which effectively treat the water surrounding a solute as a continuous medium beyond the first bound shell of water molecules. The calculated results are as precise as anybody could wish at this stage, but they do not yet describe real water.

John Maddox

Correction: Bringing photosynthesis to the bench

Since the appearance of my article commenting on a paper by J. L. Martin *et al.* in the same issue of *Nature* (363, 297 & 320–325; 1993), it has come to light that Dr Douglas C. Youvan was solely responsible for the genetic engineering of the bacteria used in the study. This would have been more evident from the original article if more explicit reference and acknowledgement had been made.

John Maddox

A stirring tale of mixing

Chris Garrett

AN understanding of mixing in the ocean is crucial for getting to grips with climate change, marine productivity and pollutant dispersal. That is why the study by Ledwell, Watson and Law, published on page 701 of this issue¹, should attract wide attention.

In recent years, a variety of somewhat indirect estimates and measurements has pointed to a smaller rate of turbulent mixing across density surfaces than used to be supposed. What Ledwell and colleagues have done is to confirm that indirect evidence by observing the spread of an artificial tracer introduced at 300 m depth in the North Atlantic. Their technically brilliant experiment also uniquely shows how the lateral spread of the tracer on a surface of equal density resembles the dispersion and mixing of a blob of cream stirred into a cup of coffee.

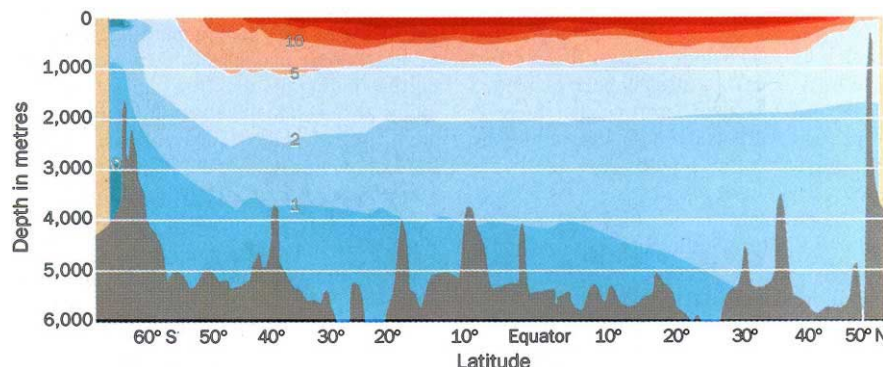
The background to this central oceanographic problem is illustrated by the figure. A simple conceptual model of the oceans has cold, dense water being formed at high latitudes by heat loss to the atmosphere, then sinking and spreading towards the Equator before upwelling in the world's oceans. The vertical profile of temperature is determined by a balance between this upwelling cold water and the downward mixing of heat from the surface, which is heated by the Sun. The estimated vertical mixing rate², K_v , of $10^{-4} \text{ m}^2 \text{ s}^{-1}$ (700 times the thermal diffusivity in the absence of turbulence) has conditioned our thinking, even though it has long been recognized that the circulation of the oceans is more complicated and that, in particular, the structure of the top kilometre or so could be determined by stirring along sloping density surfaces rather than turbulent mixing across them.

Resolving this issue is of great interest because numerical models of the ocean³ have indicated that the climatically significant transport of heat towards the poles is sensitive to the actual value of K_v , meaning that models used for climate-change predictions must specify how K_v might change — a formidable task. From theory⁴, however, it seems that if K_v is much less than $10^{-4} \text{ m}^2 \text{ s}^{-1}$ the depth of the region of rapid temperature change is largely determined by wind patterns, and the poleward transport of heat becomes less sensitive to the precise value of K_v .

Measurements of the intensity of turbulent fluctuations of temperature and velocity on scales down to a few millimetres⁵, known as 'microstructure', have rather consistently implied that, below the active surface mixed layer and down to a depth of about a kilometre, K_v is no more than $10^{-5} \text{ m}^2 \text{ s}^{-1}$. The deduc-

tion of K_v from the measurements has, however, depended on a mathematical equivalent of 'smoke implies fire' — the turbulent fluctuations could not have persisted in the face of molecular dissipation without the kind of turbulent stirring that produces vertical transport. It has been suggested, though, that smoke in one place means fire somewhere else⁶, or the other way about, and that there are sampling problems with the microstructure measurements. So a direct measurement

calculations^{8,9} and measurements^{10,11} have shown how the turbulence levels, and the K_v deduced from them, are related to the intensity of the internal waves in the ocean. These oscillatory motions in the stratified ocean interior are probably largely driven by wind at the sea surface; but they then propagate throughout the oceans and may become intensified near bottom topography, which also acts as a source of the waves as currents flow over it. So it is quite possible that K_v in the deep ocean is much bigger than $10^{-5} \text{ m}^2 \text{ s}^{-1}$, as in fact is implied by heat-budget calculations for some abyssal basins^{12,13}. Technically demanding tracer dispersion and microstructure studies for the abyssal



A section of temperature (pressure-corrected) from south to north in the middle of the Pacific Ocean suggests that cold, dense water formed near Antarctica sinks and spreads northwards, before slowly upwelling and being warmed by heat diffused down from the sea surface. (After ref. 15; contours marked are in °C.)

of vertical mixing is extremely welcome.

Following pilot experiments off the California coast, Ledwell *et al.* now report on the rate of spreading, across a density surface, of a patch of an artificial tracer (sulphur hexafluoride) that can be detected at very low concentrations. After six months the patch had spread tens of metres vertically, with an effective K_v of about $10^{-5} \text{ m}^2 \text{ s}^{-1}$, in line with the direct turbulence measurements. Subject to analysis of data obtained after a further winter (in which the mixing might be greater), this experiment thus appears to increase our confidence in deductions from turbulence measurements and, if representative, confirms that K_v in the upper ocean is small enough to be unimportant for the structure of the temperature (and salinity) fields and the poleward transport of heat. It should be stressed, however, that for prediction of climate variability on timescales of centuries, mixing rates in the deep ocean are probably important⁷, and here K_v may be larger. Deep mixing rates also have to be taken into account when dealing with deep-sea waste disposal.

By validating the microstructure approach, however, the tracer experiment has strengthened the growing framework of experiment and theory on ocean mixing rates throughout the ocean. Theoretical

ocean are still very much required, but simpler measurements of the internal wave field will provide a good first estimate of the mixing rate.

In this respect, the success of Ledwell and colleagues' experiment may have reduced the need for further such experiments. More excitingly, it shows that tracers can be used for cases where theory, and the use of microstructure techniques, are still uncertain. One widely occurring situation is where temperature and salin-

- Ledwell, J. R., Watson, A. J. & Law, C. S. *Nature* **364**, 701–703 (1993).
- Munk, W. H. *Deep-Sea Res.* **13**, 707–730 (1966).
- Bryan, F. J. *phys. Oceanogr.* **17**, 970–985 (1987).
- Welerand, P. *Phil. Trans. R. Soc. A* **270**, 415–421 (1971).
- Gregg, M. C. J. *geophys. Res.* **92**, 5249–5286 (1987).
- Holloway, G. *Dyn. Atmos. Ocean* **12**, 107–125 (1988).
- Weaver, A. J. *The Natural Variability of the Climate System on the 10–100 Year Time-Scales* (National Academy Press, Washington DC, 1993).
- McComas, C. H. & Müller, P. J. *phys. Oceanogr.* **11**, 970–986 (1981).
- Heney, F. S., Wright, J. & Flatté, S. M. J. *geophys. Res.* **91**, 8487–8495 (1986).
- Gregg, M. C. J. *geophys. Res.* **94**, 9686–9698 (1989).
- Polzin, K. L. thesis, Massachusetts Institute of Technology/Woods Hole Oceanographic Institution (1992).
- Whitehead, J. A. J. *phys. Oceanogr.* **19**, 853–861 (1989).
- Saunders, P. M. J. *phys. Oceanogr.* **17**, 632–643 (1987).
- Hamilton, J. M., Lewis, M. R. & Ruddick, B. R. J. *geophys. Res.* **94**, 2137–2145 (1989).
- Craig, H., Broecker, W. S. & Spencer, D. W. *Geosecs Pacific Expedition: Sections and Profiles Vol. 4* (US Govt Printing Office, Washington DC, 1982).

ity oppose each other in determining the vertical density structure. Subtle physical phenomena, depending on the different molecular diffusion rates of salt and heat, may then lead to different turbulent mixing rates, with the deduction from micro-structure measurements of the mixing rate of salt (or nutrients or a tracer) being particularly uncertain¹⁴. This could in fact have occurred in the present experiment but appears not to have done so, perhaps because the salinity effect on the density, while opposing that of the temperature, is weaker than in many other regions.

The experiment of Ledwell *et al.* also provides unique insight into the formation of streaks of tracer on a density surface as an initial patch is torn apart by eddies on scales of the order of 100 km or more. The narrowing of the streak is finally opposed by small-scale lateral mixing. The streaks found in this experiment were considerably wider than expected from simple two-dimensional models, implying the importance of either current shear with

height or small-scale eddies, which are often indistinguishable from internal waves but have more dispersive capability. Understanding the streakiness of a patch of tracer is more than an academic problem; it is often a central issue in pollution problems where the peak concentration of a dispersing contaminant, and not just its spatially averaged value, matters.

This and other work on the small-scale physics of the ocean are underpinned by recognition that numerical models of the ocean will never be able to resolve all the motions involved. We have to study these independently and learn how to represent them in the larger-scale models. The experiment of Ledwell *et al.* confirms that the work of the past two decades on ocean mixing is on the right track — but we still have a lot to do. □

Chris Garrett is in the School of Earth and Ocean Sciences, and the Department of Physics and Astronomy, University of Victoria, British Columbia, Canada, V8W 3P6.

PHYSIOLOGY

And brother begat nephew

John G. Vandenberg

THE notion that acquired characteristics could be inherited became fashionable in the Soviet Union after Lysenko decreed it as the genetic system for improving agricultural crops. It was false science then and remains false science now. But observations by Clark, Karpiuk and Galef on page 712 of this issue¹ suggest that, in a new incarnation, one characteristic, the sex ratio of offspring, can be acquired.

This conclusion derives from the finding that fetuses carried by mothers bearing multiple young experience an environment influenced by neighbouring fetuses. The mother and fetal neighbours affect the anatomical, physiological and behavioural development of the fetus². Clark *et al.* show that female Mongolian gerbils produce litters containing more sons if the mother underwent gestation between male fetuses than if she had female 'wombmates'. Thus, brothers can beget nephews.

Over 15 years ago, reports appeared demonstrating the surprising influence neighbouring fetuses can have on the genital anatomy of female mice and rats^{3,4}. Females gestated between two male fetuses (2M) have longer anogenital distances than females gestated between two females (2F). Females gestated between a male and a female wombmate (1M) show an intermediate condition. These reports stimulated research to explore the effect of different intrauterine positions on other systems and on other

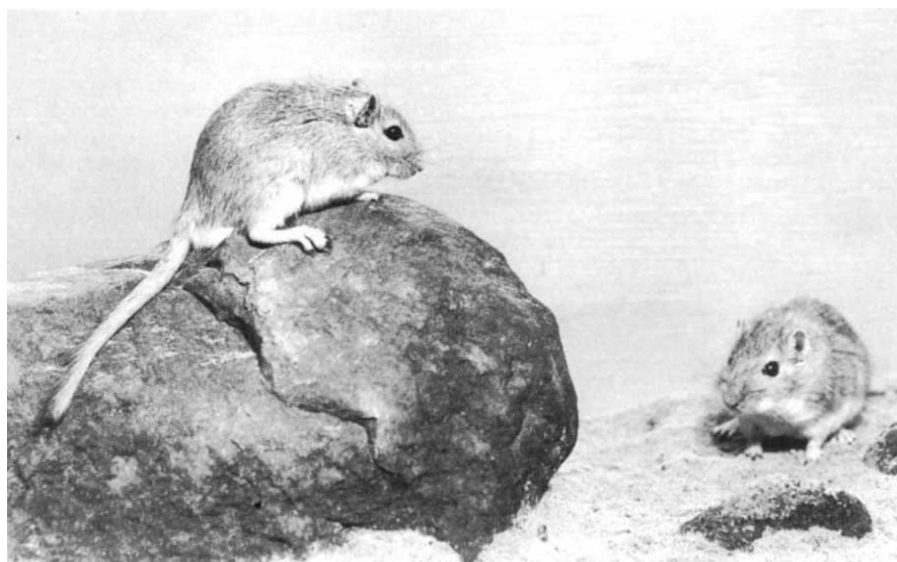
species that bear multiple young. It was quickly found that intrauterine position can influence subsequent behaviour, physiology and sexually dimorphic anatomical characteristics as well as genital anatomy². Specifically, 2M female house mice are more aggressive and less sexually active than 2F females; 2M female rats show later onset of puberty; and 2M mice have lower life-time fecundity than their 2F sisters.

Moreover, the size of the sexually dimorphic areas of the hypothalamus in

adult female rats is correlated with anogenital distance⁵, and studies of wild house mice reveal that adult females from the 2M position maintain larger home ranges than do 2F females⁶. The effects of intrauterine position are apparently due to androgens from the neighbouring males⁷. They range from brain morphology to ecological adaptation, and explain a notable source of individual variability in androgen-dependent characteristics in mammals.

Clark *et al.* now add a new life-history characteristic — sex ratio — to the list of those known to be influenced by intrauterine position. Sex ratio is important because the dynamics of a population are largely driven by the number of reproductively active females and the level of their productivity. Population models of promiscuous species typically deal only with females and assume that even a few males can provide the requisite sperm⁸. So an increase in the ratio of females can have a disproportionate effect on population dynamics. Further, an increased number of 2M females in the population due to a skewed sex ratio can increase levels of social interaction because 2M females are more aggressive and may induce increased stress in the population. Female mice experiencing social stress during pregnancy produce daughters having 2M-like genitalia regardless of intrauterine position⁹. Thus, a cascading series of events, initiated by the intrauterine position of an individual, can have profound influences on both the individual and the population of which it is a member.

Caution must be exercised in generalizing the new findings¹ to other species because intrauterine influence is especially evident in the gerbil — for example, the effects of intrauterine position on the developing male fetus are weak or nonexistent in rats and mice (the other models typically used) but strong in the gerbil¹⁰.



Mongolian gerbils — fraternally influenced.

Lumlebrook Farm, Inc.

The sensitivity of the Mongolian gerbil may make it an excellent indicator species to detect the range of these effects, but generalization to other species will require further data.

Nonetheless, the report by Clark *et al.* has heuristic value at two levels. First, a mechanism must be sought for how intrauterine position affects the sex ratio at birth. Is the female's mating behaviour and reproductive physiology altered to favour either the conception or the survival (or both) of male fetuses when she herself has been exposed to male neighbours *in utero*? Second, the findings set the stage for further research on how this system could have evolved in natural populations. The increase in the proportion of males in the population due to intrauterine influences is balanced by the disproportionately high female sex ratio in 2F females. Any factors modifying these intrauterine position effects would alter sex ratios and thereby affect population dynamics.

The old lamarkian idea, resurrected by Lysenko, of the inheritance of acquired characteristics, through environmental influences acting directly on the genome, has been thoroughly discredited. Clark *et al.* do not deal with genomic changes and thus are not tainted with this theory. Their results reveal an effect, awaiting a mechanistic and evolutionary explanation, that might be better termed the 'acquisition of an acquired characteristic' because genomic inheritance is not invoked. A male-biased sex ratio in gerbils is presumably the consequence of androgenization of the mother while she was *in utero*. As a consequence of having a higher ratio of males in her litters, her daughters would be exposed to more males *in utero* and so would produce more male-biased litters when it came to their turn to reproduce. Females would be acquiring a male-biased sex ratio from their brothers through a prenatal endocrine mechanism and, thus, their brothers would be 'enhancing the production of their own nephews by an effect on their sisters.' □

John G. Vandenberg is in the Department of Zoology, North Carolina State University, Raleigh, North Carolina 27695, USA.

1. Clark, M. M., Karpiuk, P. & Galef, B. G. Jr *Nature* **364**, 712 (1993).
2. vom Saal, F. S. J. *Anim. Sci.* **67**, 1824–1840 (1989).
3. Clemens, L. G. in *Reproductive Behavior* (eds Montagna, W. & Sadler, W. A.) 25–53 (Plenum, New York, 1974).
4. Gandelman, R., vom Saal, F. S. & Reinisch, J. M. *Nature* **266**, 722–724 (1977).
5. Faber, K. A. & Hughes, C. L. *Biol. Reprod.* **46**, 101–104 (1992).
6. Ziellinski, W. J., vom Saal, F. S. & Vandenberg, J. G. *Behav. Ecol. Sociobiol.* **30**, 185–191 (1992).
7. Clark, M. M., Crews, D. & Galef, B. G. Jr *Physiol. Behav.* **49**, 239–243 (1991).
8. Hutchinson, G. E. *An Introduction to Population Ecology* 260 (Yale Univ. Press, New York, 1978).
9. Ziellinski, W. J., Vandenberg, J. G. & Montano, M. M. *Physiol. Behav.* **49**, 117–123 (1991).
10. Clark, M. M., Bishop, A. M., vom Saal, F. S. & Galef, B. G. Jr *Physiol. Behav.* **53**, 1183–1187 (1993).

Nonlinearity and chaos at work

Henry D. I. Abarbanel

CHAOS is moving out of the mathematics department and into engineering design. That was the take-home message to emerge from a symposium* held last month. The possibilities offered by the vast expansion of knowledge about nonlinear systems, and their ubiquitous chaotic orbits, were made plain in reports of new directions in design of a variety of technologies.

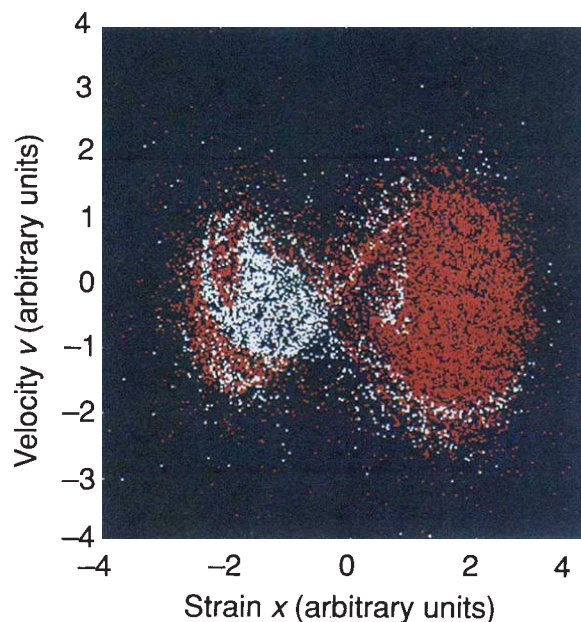
Engineering aims to bring ideas out of the laboratory and onto the production

descriptions and provide new insight into useful, bounded oscillations of direct engineering interest. As we systematically map them out, bifurcations change from design threats to design opportunities.

Similarly with chaos — the very word puts off the practical engineer. But if we think nonlinearly we see chaos as the chance to explore a broad set of possible states of the engineering system in a deterministic and useful fashion. Although it is indistinguishable, in a practical sense, from 'random' motion when seen traditionally, when viewed in the multidimensional space it properly requires, chaos has exploitable and useful regularities that can be captured in 'rules' for engineering purposes.

The ideas of bifurcation theory have been incorporated in automated software (R. H. B. Fey, Technical University Delft) and used to discover the domains of regular and irregular motions available to driven structural beams with amplitude-limiting flanges. Despite the large number of vibrational modes allowed to the beam, its behaviour is quite low-dimensional and each feature of its observed motion is well described by a few degrees of freedom. The verification of the low-dimensional behaviour by this kind of experiment underlines the engineering design potential of these

tools for structural applications. Rolling-stock design and railroad bed properties would seem to have been fully exploited by the practical engineers of the nineteenth century. But a haunting reminder of how little we know was provided in the description of the 'hunting' motion of railway vehicles (H. True, Technical University of Denmark, Lyngby). Hunting is a vibration transverse to the tracks which occurs at critical velocities in all railroad operations; it deforms and destroys the roadbed, is a severe maintenance and safety problem, and limits train speeds. True exposed this motion as an integral property of a low-order nonlinear system describing the contact between the wheels of the train and the track. Hunting, which is asymmetric between the right and left track, is a standard feature of the low-order system, and it arises simply as part



The basins of attraction for a stiff beam hanging in a double well potential produced by local magnets. There are two fixed points of the motion, leaning left or leaning right. The phase plane (initial position x versus initial velocity v) is marked with red points for which the beam moves in time to the reclining right position. For points marked in white, the beam ends up leaning left. (Figure courtesy of J. P. Cusumano.)

floor, and the design process in this transfer of technology has traditionally rested on linearized dynamical equations or linearizations about 'operating points'. It was this tradition that came under challenge at the symposium.

Two broad topics, bifurcations and chaos, underlay many of the talks. The instabilities occurring at bifurcation points (where one characteristic motion of a system splits into new motions as a physical parameter is varied) have always constituted warning notices of dangerous territory for engineers. When one thinks nonlinearly, however, these become routes for discovering how real, certainly nonlinear, systems cure the explosive behaviour predicted by solely linear

*Nonlinearity and Chaos in Engineering Dynamics, IUTAM Symposium, University College London, UK, 19–23 July 1993.

of the universal bifurcation sequence of such models. The design of practical controls of this and other hunting in the context of small nonlinear models now becomes possible. Such controls could save some US long-haul freight carriers a great deal of money, for at present they must limit train speeds and build facilities to 'turn around' long trains (which can be two kilometres and more in length) to avoid asymmetric wear.

Fractal and chaotic behaviour of surfaces cut by machine tools (F. C. Moon Jr, Cornell University) were revealed as simple geometric objects when seen in the appropriate state space. Moon's laboratory observations of machine-tooled surfaces allow the engineer to classify the surfaces cut from various materials. The fractal dimension of the surfaces was roughly 1.85 and is directly connected with the wavenumber spectrum of the surface. Most important for engineering considerations is the discovery that simple time-delay equations of the form $x'' + Ax' + Bx = x(t) - x(t - T)$ can account for the experiments. Further, Moon emphasized the importance of the experimentally derived probability distributions for the data in these deterministic systems — this too is a helpful design tool.

Experiments were the keynote of the symposium (and here I should say I do scant justice to the theory discussed). A vibration absorber for rotating machinery (S. W. Shaw, University of Michigan) achieves its goals of removing oscillations, induced by external torques, by prescribing paths for counter-rotating dynamical elements based on an understanding of the subharmonic bifurcation properties of the whole system. These subharmonic motions precisely cancel rotation-induced vibrations in helicopter and automotive machinery. The use of the classical period-doubling route to chaos for the design of a vibration absorber was impressive. We can expect these nonlinear absorbers to appear in products of the Ford motor company (Ford sponsored the work).

Further experiments focused on other simple systems. The oscillations of a stiff beam with two equilibrium states (J. P. Cusumano, Penn State University) gave an experimental demonstration of the fractal boundary of the basin of attraction between the equilibria. Engineers (indeed, most of us) wish to know where mechanical systems will go after we start them. Cusumano observes (see figure) that in as simple a system as his beam there are regions of state space where the answer to this question may sensitively depend on where you start. The parts of phase space (x versus v in the figure) coded with white dots end with the beam leaning left, whereas red dots end with the beam reclining right. This work highlights the practical aspects of identifying the boundary (fractal) between these choices

and provides the basis for precise design to achieve the desired target in application.

The import of these sort of experiments is, first, that complex behaviour in elementary mechanical oscillators can be accurately described by simple nonlinear models. Second (and even more to the point for engineering design), low-dimensional descriptions of complex observed behaviour in engineering mechanical systems composed of nonlinear elements can be realized and thus exploited.

As Sir James Lighthill pointed out in his opening address, for more than 300 hundred years — ever since the appearance of

Newton's *Principia* in 1686 — we have been restricted to analysing the regular behaviour of mechanical systems. We are now liberated from the restrictive legacy of Newton, Lagrange and Hamilton, where we could dwell only in limited phase space. The fruits of the work of the past decade mean that we have moved from being able to achieve chaos by design to being able to achieve design by chaos. □

Henry D. I. Abarbanel is at the Scripps Institution of Oceanography and the Institute for Nonlinear Science, University of California San Diego, La Jolla, California 92093, USA.

MUSCULAR DYSTROPHY

Muscling in on gene therapy

Helen M. Blau

SINCE the cloning of the dystrophin gene, one of the largest isolated to date^{1,2}, Duchenne muscular dystrophy (DMD) has been of great interest as a target for gene therapy. The disease afflicts one in 3,500 boys each year and is the most severe of the muscular dystrophies, causing progressive muscular wasting and death by about age 20. Although the cloning of the gene has revolutionized diagnosis, the only treatment available to DMD patients is palliative. Hence the attraction of gene replacement as a potential therapy.

The paper by Jeffrey Chamberlain and colleagues (Cox *et al.*) on page 725 of this issue³ is a significant step forward. This report provides evidence that expression of a recombinant gene, encoding the missing protein dystrophin, can prevent muscle destruction and preserve muscle function in an animal model of muscular dystrophy. It shows that in *mdx* mice (which, like DMD patients, have an aberrant dystrophin gene) over-expression of a full length dystrophin complementary DNA not only leads to normal muscle fibre morphology, but also to normal force and power in the diaphragm; this latter finding is of particular interest as respiratory failure is a leading cause of death in man.

A major surprise is that muscle is so forgiving with respect to the amount of the dystrophin molecule. In the transgenic mice created by Cox *et al.* it is present at 50 times normal levels without adverse effects. Ordinarily dystrophin makes up only 0.002 per cent of total skeletal muscle protein. Such low levels may be ensured by the 2,300-kilobase size of the gene from which a single transcript is probably produced each day. Disease severity appears to depend in part on the quantity of dystrophin, because little or no protein results in DMD, whereas intermediate

levels are associated with more benign Becker muscular dystrophy⁴. If the findings of Cox *et al.* in the mouse also relate to man, there may be no such thing as too much dystrophin. That would be highly encouraging, for it is not yet possible to regulate precisely levels of expression of recombinant genes *in vivo*.

Given the presumed structural function of dystrophin, the large size of the protein (M_r 427,000 (427K)) might be expected to be critical to its function. However, individuals who harbour either a deletion of the gene that eliminates one-half of the dystrophin molecule, or a duplication of the gene that results in a 50 per cent increase in protein size, may have the milder Becker muscular dystrophy^{5,6}. Moreover, large deletions often lead to a less debilitating disease than small deletions, if they do not disrupt the reading frame⁷. Apparently, truncated or enlarged dystrophin is better than no dystrophin, if it contains critical domains such as the carboxy terminus. Dystrophin shares homology with the cytoskeletal proteins, α -actinin and spectrin⁸, is localized at the plasmalemma⁹, and is required to anchor certain integral membrane proteins, such as dystrophin-associated glycoproteins (DAGs)¹⁰, which may control calcium flux. Evidence that DAGs may be important to muscle function comes from patients with a heritable muscular dystrophy, in which dystrophin is normal but a DAG of 50K is deficient¹¹. In mice, certain DAGs are reduced in the absence of dystrophin, and restored in its presence^{3,12}; but whether such restoration can also be achieved upon expression of a truncated dystrophin molecule remains to be determined.

The delivery of dystrophin poses a unique challenge because the gene must be expressed in all muscles. Cox *et al.* met this challenge by injecting the recom-

Palaeontology and existential angst

DANISH philosopher Søren Kierkegaard (1813–1855) receives an unusual posthumous honour on page 709 of this week's issue: Graham Budd has named a fossil after him. The excuse, that both philosopher and fossil are associated with Copenhagen, is prosaic. But the fact that *Kerygmachela kierkegaardi* is one of those sanity-challenging early Cambrian fossils, so vividly described by Stephen Jay Gould in *Wonderful Life*, is perhaps appropriate in the light of Kierkegaard's revolt against objective reality. What would have amused the Dane somewhat less is that, strange as they are, the unusual arthropods from the Cambrian Explosion may inform objective schemes that resolve pressing problems of arthropod phylogeny. Budd interprets *Kerygmachela* as a lobopod. Known today from the tropical, terrestrial onychophorans such as *Peripatus*, lobopods seem to have been a highly diverse group in early Cambrian times, ranging from the *Peripatus*-like *Aysheaia* to *Hallucigenia*, so much the *doyen* of weird Cambrian fossils that it has inspired the creation of science-fictional aliens. The lobopods are characterized by their



stumpy, unsegmented limbs, and *Kerygmachela* seems to have an abundance of these. But it also has large, lateral exten-

sions of the body that Budd interprets as gills. These resemble two other structures in Cambrian oddities, *Opabinia* and *Anomalocaris*, Cambrian top predator, and — at first sight — an unlikely member of the Lobopodia. *Opabinia* and *Anomalocaris* are not known to have lobopod-like limbs, but that doesn't mean that they aren't there. Diligent examination revealed the requisite organs in *Hallucigenia*. But there is no doubt that *Kerygmachela* has both, and Budd takes the bold step of marrying them as two components of a prototype biramous limb, as found in crustaceans, chelicerates and trilobites: but not insects, myriapods or modern onychophorans such as *Peripatus*. The phylogenetic consequences of this are profound, because it means that the Lobopodia is not a natural group allied to uniramous arthropods, but a grab-bag containing both uniramous and biramous forms, from which the different groups of arthropod emerged separately. If this scheme is accepted (a moot point given the

continuing debate about arthropod phylogeny), a welcome order may emerge amid the Cambrian chaos. Henry Gee

binant gene into a fertilized mouse egg, an option not available in humans. Direct injection of DNA into cardiac and skeletal muscles has led to gene expression, but the number of cells that take up the DNA is so low that this route of delivery, although useful for some disease states, will not yet suffice for the treatment of DMD^{13,14}. Well-characterized viral vectors cannot accommodate a 14-kb cDNA, though they could be used to deliver a 7-kb 'minigene', the size of that found in some mildly affected patients with deletions⁵. However, the safest replication-defective retroviral vectors will not infect differentiated muscle fibres, which are postmitotic. Adenoviral vectors can gain access to non-dividing cells and lead to dystrophin expression in up to 50 per cent of muscle fibres, but unless the vectors are introduced into mice at a neonatal stage¹⁵, gene expression is only transient, a finding which may limit their general usefulness. Another drawback of adenoviral vectors is that they currently lead to widespread gene expression in several tissues. One way to improve their potential would be to make them tissue-specific like the dystrophin transgene reported by Cox *et al.*, which is restricted

in its expression to muscle and brain by a muscle creatine kinase enhancer.

Another concern is which gene construct to deliver. The dystrophin transcript is not only large, but also alternatively spliced and initiated from several different promoters^{16,17}. The functions of the different transcripts, and the large and small proteins they produce in fetal and adult development, and in different tissues such as the brain, remain unclear. Patients suffering from DMD not only have impaired muscle function but also impaired cognitive faculties. Use of the minigene that can be accommodated in viral vectors did not meet with the same degree of success in transgenic animals as the full length cDNA reported by Cox *et al.*, although it remains unclear whether this was due to the construct or to the low level of expression achieved in these experiments¹⁸. It could be argued that the mild disease manifested in patients harbouring the minigene constitutes evidence that a truncated dystrophin is therapeutic. But the minigene is only known to occur in a few pedigrees, and the possibility remains that the genetic background, or other proteins present in these individuals' muscles, might be responsible for

their mild symptoms. That such proteins exist is suggested by studies of the *mdx* mouse; unlike man, this animal is only transiently affected by the lack of dystrophin it experiences throughout life¹⁹. So a major question is whether the beneficial effects reported by Cox *et al.* in the *mdx* mouse will also pertain to DMD patients.

An alternative approach to gene replacement is myoblast transplantation. This has the advantage that the entire dystrophin gene can be delivered with its regulatory sequences intact, thereby ensuring that different transcripts and proteins will be produced at the appropriate times. The approach was stimulated by animal studies which indicated that injected myoblasts cross basal lamina and become an integral part of mature host muscle fibres^{20,21}. Phase I clinical trials revealed no deleterious effects of myoblast transfer to the muscles of DMD patients, and in some cases detected dystrophin gene expression^{22–24}. However, the number of fibres expressing dystrophin in muscles of DMD patients was much lower than that observed previously in *mdx* mice, possibly because the disease is more severe in man, and had progressed

too far, or because of technical problems associated with the delivery of cells into human muscles. To distinguish between these possibilities, we need a method for marking and tracking injected myogenic cells in humans, and studies of human muscles that are not disrupted by the DMD disease process. Although myoblast transfer is unlikely to constitute a cure, because brain function will not be altered and all muscles including the heart and diaphragm would need to be separately injected, it might improve the quality of life for DMD patients by allowing them to be mobile for a longer period.

Genetic engineering strategies other than gene replacement can be envisaged that could compensate for the lack of dystrophin in DMD. One is to upregulate genes on autosomes encoding dystrophin-like molecules, such as utrophin, that are normally expressed only early in development but, like dystrophin, colocalize with DAGs at the neuromuscular junction^{25,26}. Another might be to induce alternative patterns of dystrophin RNA splicing, which change the reading frame and yield functional short proteins. A better understanding of the inflammatory response, the nature and role of calcium channels, and other mechanisms that lead to muscle degeneration and fibre death in DMD could be exploited to preserve muscle-fibre function in the absence of dystrophin. But although these approaches are exciting, they are not likely to be realized for some time. Gene replacement, for all the caveats, is the strategy for the near future. □

Helen M. Blau is in the Department of Molecular Pharmacology, Stanford University School of Medicine, Stanford, California 94305, USA.

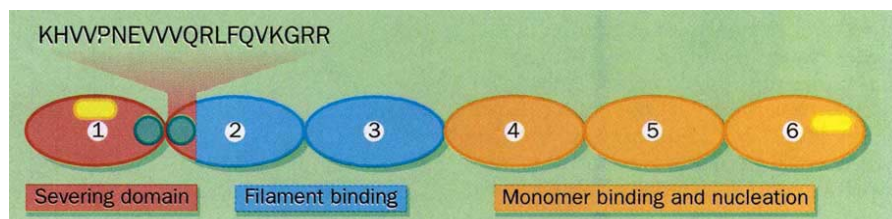
A slice of the actin

Paul Janmey

THE ability of gelsolin and some of its homologues to break actin filaments is a unique biochemical reaction. With the report by McLaughlin and colleagues on page 685 of this issue¹ we get a better idea of how that reaction works, for the authors provide the first structural model of the process.

The long slender filaments of actin transmit tension in muscles, participate in the machinery of cell protrusion, and act as scaffolds to which a host of other proteins bind. They are remarkably stable

further addition or release of subunits. Gelsolin's ability to accomplish these feats stems from its having six similar domains of M_r 14,000, all of which might bind actin. Intact gelsolin binds only two actin monomers, so different domains compete with each other or bind to different sites on actin to achieve the variety of possible effects. The filament-severing activity resides within the first two domains, because segment 1 with a few additional residues from segment 2 has nearly the same severing efficiency as the intact molecule.



Multidomain structure of gelsolin. Biochemical studies have identified actin monomer binding activities in segments 1 and 4–6, and filament side-binding activity in segments 2–3. Domains 2–6 are sufficient for nucleation, and segments 1–3 for severing. The smallest fragment capable of severing actin filaments is segment 1 plus the 20 residue peptide denoted at the amino terminus of domain 2. Green circles indicate sites where polyphosphoinositides bind, yellow areas are Ca^{2+} -binding or Ca^{2+} -sensitive sites.

polymers that can contain several thousand non-covalently bound subunits, but break at a negligible rate. Although able to withstand mechanical forces and the binding of dozens of different proteins, actin filaments are exquisitely sensitive to gelsolin, which severs them with almost perfect efficiency. This cleavage is strictly regulated. Micromolar concentrations of Ca^{2+} or low pH activate gelsolin to sever, whereas certain polyphosphoinositides inhibit severing and in some cases release gelsolin and allow filament growth². The result is a reversible system, which regulates polymer length and therefore the viscoelasticity of the intracellular polymer network in an optimally efficient way that synthetic polymer chemists can only dream about. The structure reported by McLaughlin *et al.* (see figure) suggests how both aspects of this regulation occur.

Gelsolin is so-called because it was first identified as the factor responsible for liquefying (solating) gels formed by macrophage extracts³. A secreted form⁴ encoded by the same gene⁵ is part of a scavenger system to clear actin filaments from the vasculature⁶. Gelsolin affects actin polymerization in a number of ways. As well as severing filaments, it binds actin monomers and promotes polymerization by accelerating the slow nucleation step. Whether it nucleates a new filament or cuts an old one, gelsolin remains bound to the fast-exchanging filament end (the barbed end) preventing

It is in this context that the crystal structure of gelsolin segment 1 in a complex with monomeric actin provides part of the picture of how gelsolin can pry apart the bonds holding the actin filament together.

Segment 1 is composed of a four-stranded β -sheet with α -helices bound to each side. Sequence analysis predicts that each of the six gelsolin domains has a similar structure. One of the helices and residues from the corresponding side of the sheet in segment 1 bind into the cleft between sub-domains 1 and 3 of the actin structure⁷. An interesting feature of this contact is a core of hydrophobic residues encircled by a ring of hydrogen bonds joining polar residues. By locating the site of contact on the model of F-actin, McLaughlin *et al.* conclude that gelsolin segment 1 binds tangentially to the actin filament, making a glancing slice between a pair of actin subunits in one of the two long-pitch helices of F-actin. They also conclude that disrupting one of the four actin-actin contacts that each subunit forms is enough to break the filament. This conclusion is somewhat indirect, in that segment 1 by itself has little or no severing activity and relies on the presence of the second bond to actin provided by segment 2. The simplest interpretation is that segment 2 merely helps segment 1 find its target, but the additional binding site may play a more active role in severing.

The model of the gelsolin-actin com-

1. Koenig, M. *et al.* *Cell* **50**, 509–517 (1987).
2. Burghes, A. H. M. *et al.* *Nature* **328**, 434–437 (1987).
3. Cox, G. A. *et al.* *Nature* **364**, 725–729 (1993).
4. Hoffman, E. P. *et al.* *New Engl. J. Med.* **318**, 1363 (1988).
5. England, S. B. *et al.* *Nature* **343**, 180–182 (1990).
6. Angelini, C., Beggs, A. H., Hoffman, E. P., Fanin, M., & Kunkel, L. M. *Neurology* **40**, 808–812 (1990).
7. Liechti-Gallati, S. *et al.* *Hum. Genet.* **81**, 343–348 (1989).
8. Davison, M. D. & Critchley, D. R. *Cell* **52**, 159–160 (1988).
9. Arahata, K. *et al.* *Nature* **333**, 861–863 (1988).
10. Campbell, K. P. & Kahl, S. D. *Nature* **338**, 259 (1989).
11. Matsumura, K. *et al.* *Nature* **359**, 320–322 (1992).
12. Matsumura, K., Lee, C. C., Caskey, C. T. & Campbell, K. P. *FEBS Lett.* **320**, 276–280 (1993).
13. Wolff, J. A. *et al.* *Science* **247**, 1465–1468 (1990).
14. Acsadi, G. *et al.* *Nature* **352**, 815–818 (1991).
15. Ragot, T. *et al.* *Nature* **361**, 647–650 (1993).
16. Feener, C. A., Koenig, M. & Kunkel, L. M. *Nature* **338**, 509–511 (1989).
17. Nudel, U. *et al.* *Nature* **337**, 76–78 (1989).
18. Wells, D. J. *et al.* *Hum. molec. Genet.* **1**, 35–40 (1992).
19. Bulfield, G., Siller, W. G., Wight, P. A. L. & Moore, K. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1189–1192 (1984).
20. Partridge, T. A., Morgan, J. E., Coulton, G. R., Hoffman, E. P. & Kunkel, L. M. *Nature* **337**, 176–179 (1989).
21. Blau, H. M., Dhawan, J. & Pavlath, G. K. *Trends Genet.* **9**, 269–274 (1993).
22. Gussone, E. *et al.* *Nature* **356**, 435–438 (1992).
23. Huard, J. *et al.* *Muscle & Nerve* **15**, 550–560 (1992).
24. Karpati, G. *et al.* *Ann. Neurol.* **34**, 8–17 (1993).
25. Matsumura, K. *et al.* *Nature* **360**, 588–591 (1992).
26. Tinsley, J. M. *et al.* *Nature* **360**, 591–593 (1992).

plex contains a few surprises. One is that segment 1 traps a Ca^{2+} ion in a pocket which lacks a seventh coordination site for Ca^{2+} and which resembles a site in annexins. The site surrounding the trapped Ca^{2+} in segment 1 is proposed to be a binding site for polyphosphoinositides, which is consistent with studies showing that deletions in this region of segment 1 render its binding to actin insensitive to polyphosphoinositides. Another surprising feature is a Ca^{2+} bound at the interface by both actin and gelsolin. This very unusual sharing of bonds suggests a new means by which Ca^{2+} can regulate protein-protein interactions, although *in vitro* the binding of segment 1 to actin appears to be Ca^{2+} -insensitive.

Among the questions that this structure raises are why only some gelsolin homologues sever filaments, and why point mutations outside domain 1 selectively disrupt severing but not other actin-binding functions⁸. For instance, the *Dictyostelium* protein severin and the macrophage protein gCap39 or MCP have structures similar to gelsolin domains 1–3 and are regulated by Ca^{2+} and phosphoinositides. But whereas severin cuts F-actin⁹, gCap39/MCP merely binds the end of the filament. The structure of myosin S1 and its complex with actin, published last month¹⁰, shows that myosin heads also bind to a site on actin very near where gelsolin binds, but myosin presumably moves along rather than breaks the filament.

One remaining question is how the various segments of intact gelsolin work together to sever the filament. McLaughlin *et al.* propose that segment 1 and its closest relative segment 4 wrap around the filament to snip equivalent bonds on each of the long-pitch helices, although an active destabilization of the filament by segment 2 or 3 also remains a possibility. A decision between these alternatives awaits the solution of the entire gelsolin crystal structure. Whether gelsolin is a Ca^{2+} -activated switchblade or a pair of scissors is not yet clear, but at least one of the blades is now revealed. □

Paul Janmey is in the Division of Experimental Medicine, Brigham and Women's Hospital, 221 Longwood Avenue, Boston, Massachusetts 02115, USA.

The fall of the C–C bond

William D. Jones

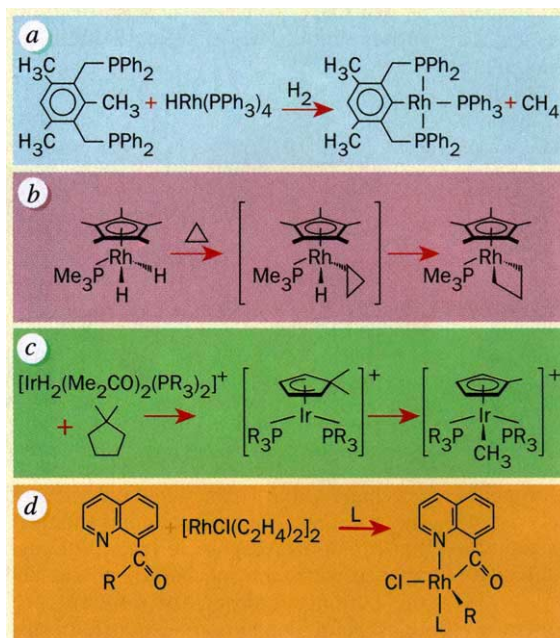
THE petroleum industry takes large hydrocarbon molecules and cleaves them into smaller pieces by the addition, at high temperature, of hydrogen in the presence of a heterogeneous catalyst. One holy grail of the modern chemist is to be able to cleave carbon–carbon bonds in unreactive saturated hydrocarbons and aromatics in a selective fashion. As described on page 699 of this issue¹, David Milstein and

ies for breaking down these hydrocarbons are generally limited to obtaining a range of cleavage products from the high temperatures and extreme conditions used — the hydrocarbons are heated to temperatures at which bond homolysis occurs, with the heterogeneous surface mediating the subsequent chemical reactions to give smaller hydrocarbon fragments. The drawback is not only that high temperatures are required, but that the process is rather unselective.

Milstein and co-workers now show that C–C bonds between an aromatic ring and a methyl group can be selectively cleaved under mild conditions. The importance of this discovery can be put into context by considering the previous examples of C–C bond cleavage by homogeneous transition-metal complexes (see figure). Several examples have been found in which the C–C bond cleaved is part of a strained ring system, with this strain providing the thermodynamic driving force for the cleavage reaction; cyclopropanes and cyclobutanes dominate this class of substrate². A second class of molecule in which C–C cleavage has been seen involves the formation of an aromatic cyclopentadienyl ring by migration of an alkyl group from the C_5 ring to a metal centre³. Less common are examples in which the C–C bond adjacent to a pendant ketone is cleaved to give a metal acyl complex⁴.

The cleavage of C–C bonds is believed to be difficult in that metal centres that can insert into that bond can also insert into a C–H bond. The latter surround the periphery of the hydrocarbon molecule, so many reactive metals are seen to insert exclusively into the C–H bonds of alkanes and arenes.

What Milstein and colleagues have done is to make use of the coordination effect to enhance the reactivity of a methyl–phenyl bond. A notable feature of this system is that C–H bond activation occurs readily, and is much more rapid than the C–C cleavage reaction. Ultimately, however, C–C cleavage does occur, and Milstein's group used an excess of hydrogen to remove one of the hydro-



Examples of C–C bond cleavage reactions. *a*, Milstein's selective methyl cleavage with methane loss (Ph, phenyl). Only the methyl group between the pendant phosphine groups is removed. Also note that the cleavage does not relieve strain in the substrate. The rhodium complex also loses three PPh_3 ligands in the reaction (not indicated). *b*, A C–C cleavage relying on relief of strain in cyclopropane, by Periana and Bergman² (Me, methyl). *c*, A methyl migration C–C cleavage that relies on formulation of an aromatic $\text{C}_5\text{H}_4\text{CH}_3$ ligand, by Crabtree and Dion³ (R, *p*- $\text{C}_6\text{H}_4\text{F}$). *d*, An activated C–C cleavage in a coordinating ketone, by Suggs and Jun⁴.

colleagues have made success in that quest more likely — they have found a homogeneous transition-metal complex that selectively cleaves one of three methyl groups attached to an aromatic ring.

The ability of organic chemists to synthesize complex molecules lies mainly in forming carbon–carbon bonds selectively and sequentially. Methods for the cleavage of carbon–carbon bonds are far more limited, and require the presence of functional groups or unsaturation (for example ozonolysis of a C–C double bond). Saturated long-chain hydrocarbons are virtually unreactive to most common organic reagents, and as a consequence their chemistry has not been developed. The processes used in petroleum refiner-

- McLaughlin, P. J., Gooch, J. T., Mannherz, H.-G. & Weeds, A. G. *Nature* **364**, 685–692 (1993).
- Yin, H. L. *Adv. exp. Med. Biol.* **255**, 315–323 (1989).
- Yin, H. L. & Stossel, T. P. *Nature* **281**, 583–586 (1979).
- Chaponnier, C., Borgio, R., Rungger-Brandt, E., Weil, R. & Gabbiani, G. *Experientia* (Basel) **35**, 1039–1040 (1979).
- Kwiatkowski, D. J. *et al.* *Nature* **323**, 455–458 (1986).
- Lee, W. & Galbraith, R. *New Engl. J. Med.* **326**, 1335–1341 (1992).
- Holmes, K., Popp, D., Gebhard, W. & Kabsch, W. *Nature* **347**, 44–49 (1990).
- de Aruda, M., Bazari, H., Wallek, M. & Matsudaira, P. *J. Biol. Chem.* **267**, 13079–13085 (1992).
- Eichinger, L. & Schleicher, M. *Biochemistry* **31**, 4779–4787 (1992).
- Rayment, I. *et al.* *Science* **261**, 58–65 (1993).

RÉSUMÉ

carbon fragments irreversibly as methane, thereby driving the C–C cleavage reaction to completion. The organometallic product also forms a strong metal–phenyl bond, which undoubtedly contributes to making the C–C cleavage reaction more favourable.

One reason for the difficulty of inserting into the C–C bonds in saturated hydrocarbons is that the two carbon atoms contain 'directed' sp^3 hybridized orbitals that lie directly along the bond axis. The inaccessibility of these orbitals makes the bond less reactive. Milstein's system cleaves the bond between sp^2 and sp^3 hybridized carbons, which are abundant in petroleum resources. The sp^3 hybridized carbon contains additional orbitals with which the metal may interact, allowing it to get close to the C–C bond before the actual cleavage.

That C–H bond activation occurs reversibly, but that C–C activation once it takes place is made irreversible by

hydrogenation implies that the anchoring of the aromatic group to the metal centre may not be a requirement for C–C cleavage to occur. If so, then free alkylated aromatic hydrocarbons might also be subject to attack. Many metal systems are known to reversibly activate aliphatic hydrocarbons, so it may ultimately become possible to cleave these sp^3 – sp^3 C–C bonds. In this sense, the new work opens the door to further studies that may change the way petroleum is processed. □

William D. Jones is in the Department of Chemistry, University of Rochester, Rochester, New York 14627, USA.

1. Gozin, M., Weisman, A., Ben-David, Y. & Milstein, D. *Nature* **364**, 699–701 (1993).
2. Periana, R. A. & Bergman, R. G. *J. Am. chem. Soc.* **108**, 7346–7355 (1986).
3. Crabtree, R. H. & Dion, R. P. *J. chem. Soc., chem. Commun.* 1260–1261 (1984).
4. Suggs, J. W. & Jun, C.-H. *J. Am. chem. Soc.* **106**, 3054–3056 (1984).

ASTRONOMY

Being around at the death

Harm Habing and Paul Murdin

As its hydrogen fuel runs out, in five billion years, the Sun will become a red giant — a star so large (30 per cent of the size of Jupiter's orbit) that the Earth will be swallowed and evaporated. At its centre it will have a small, very hot and dense core, made of oxygen and carbon; the core will be surrounded by an envelope of helium and hydrogen at very low density. This envelope will be gradually ejected into space. Jupiter will raise large tides on the solar surface, and make the Sun spin faster, causing the solar wind to blow into an elliptical shell. Finally, when the solar core is bare, the Sun will become a white dwarf — a tiny, hot star that slowly fades and cools. The white dwarf will ionize the solar wind, and other civilizations in the Galaxy will observe a fine non-circular planetary nebula where the Sun once was. Such is our future, as foretold by N. Soker (Oranim, Israel) at a meeting* last month.

The matter that surrounds many dying stars was the theme for the conference. It was the first time that astronomers interested in modest stars, such as our Sun, met colleagues with an interest in more massive stars, which, in contrast to our Sun, die as supernovae. In both cases, circumstellar material alters the appearance and the evolution of the object.

The detailed study of this material has become possible in recent times thanks to

technical progress and luck. Luck is represented by two nearby supernovae — SN1987A in the Large Magellanic Cloud, and another this year (SN1993J) in the nearby spiral galaxy, M81 — both of which revealed circumstellar material generated earlier by the exploding stars. Technical progress has come in the form of the sensational maps taken with the new French–German millimetre-wave array of interferometers (see figure on next page). M. Guelin (IRAM, Grenoble) showed these maps of circumstellar nebulae with an angular resolution of about 100 stellar radii, maps that allow the measurements of temperatures, densities and outflow velocities.

What controls the outflow from red giant stars? Outflowing matter cools rapidly and a fraction of the gas condenses into small solid particles. The solid particles are accelerated outwards by radiation pressure from the star, and they carry the gas along by friction. The warm particles are detected by their infrared emission. They exist, but why? Normal condensation processes in operation on Earth do not work so far out of thermodynamic equilibrium.

For one of the two main species — carbon-containing particles — the chemistry, if not the physics, is available from the 'Detroit connection': the study of the formation of soot in the exhaust of cars (A. Tielens, NASA/Ames). Acetylene (C_2H_2) combines to form benzene rings, which agglomerate into planes of polyaromatic hydrocarbons. These stack into

Heart of coke?

WHAT is at the centre of the Earth? Going by the cosmic abundances of the elements, the core is chiefly iron with about 8% nickel, but a lighter element must also be present to account for its density. Sulphur is the popular candidate; carbon has been reckoned too volatile to be present in substantial concentrations. But Bernard Wood believes that something has been overlooked: the volatility strongly depends on pressure (*Earth planet. Sci. Lett.* **117**, 593–607; 1993). From calculated phase diagrams and experiments, he shows that at 0.01–5 GPa, pressures applicable during the Earth's accretion, 2–4 wt% carbon could have been stable in the iron-rich melt. Even if the core started out overall with a mere 0.5–0.6 wt% carbon, the phase that first precipitated out, and hence forms the solid inner core, could well be iron carbide (Fe_3C), rather than an iron–nickel solid solution.

End game

THE *Listeria bacillus*, once infamously said to nestle in every Camembert, drives itself by throwing out an ever-growing bundle of actin filaments from one end. Its passage through the host cell cytoplasm is marked by an exhaust of actin. Last year C. Kocks *et al.* reported that the formation of the actin tail fibre was inseparable from the expression of a gene, coding for a protein which they called ActA. They have now shown by immunofluorescence (*J. Cell Sci.* **105**, 699–710; 1993) that the protein is indeed absent from one end of the bacterial surface and abundant at the other. It is laid down asymmetrically at cell division, and evidently seeds actin polymerization. Another protein, not yet identified, must cross-link the F-actin filaments into a tail. Kocks *et al.* offer the conjecture that actin nucleation is regulated by the local ATP/ADP balance and note that in the related pathogen, *Shigella*, a protein which may do the same office as ActA possesses ATPase activity.

Soldering on

LEAD-free solders are highly desirable, given the toxicity of lead. Alloys of silver and tin offer high resistance to fatigue — welcome news to those who torture their circuits by repeated bending or thermal cycling — but frequently suffer from non-uniformity and surface roughness. By adding just a smattering of zinc to an alloy of Sn–3.5 wt% Ag, M. McCormack and colleagues have now developed a solder in which the microstructure is much more uniform. Gone are the pervading dendrites of nearly pure, soft tin found in the binary alloy; instead, an alloy with 1 wt% added zinc shows a fine dispersion of Ag_3Sn precipitates, giving it a peak strength 48% higher than the zinc-free equivalent, and an order of magnitude better high-temperature creep resistance.

*Circumstellar Media in the Late Stages of Stellar Evolution, 34th Herstmonceux Conference, Royal Greenwich Observatory and the Institute of Astronomy, Cambridge, UK, 12–17 July 1993.

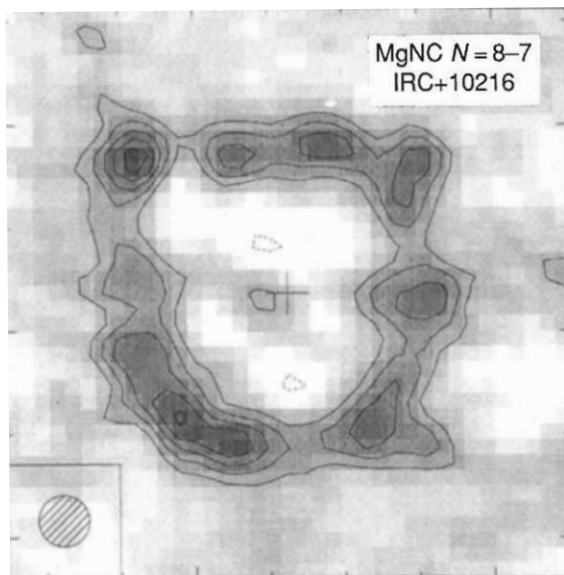
platelets and coagulate into soot particles or curl into fullerenes. The origin of the other kind of solid particles, silicate material, remains as unclear as it always has been.

Light pressure by itself cannot accelerate the gas to escape speed, and additional dynamic processes in the lower atmosphere are needed to explain the phenomenon. We have only the glimmerings of an understanding of what these processes could be. An added difficulty is that the flows of gases from and around the stars are often very clumpy (J. Dyson, Manchester Univ.); large separate pieces of dense material embedded in a smoother background affect the behaviour and appearance of the overall flow. The origin of these clumps is not understood, which brings some fundamental uncertainty into computer simulations of gas flows — just at the point when computer simulation of two-dimensional flows has become a well-established tool. Astronomical life would be easier without clumps but the clumps do seem to exist in planetary nebulae: molecules, by contrast to ions or atoms, have been found at their edges. The molecules ought to have been dissociated by strong, local ultraviolet radiation. But if the molecules are inside dense clumps, they might be able to survive (P. Huggins, New York Univ.).

The large variety in appearance of planetary nebulae, showpieces for astronomical imaging (H. Schwarz, ESO Chile), is a result of only a few basic forms seen at different angles. These forms are a consequence of multiple, consecutive outflows of gases. For example, axisymmetric flows from the white dwarf star bumping into spherically symmetric flows from the earlier red giant stage cause focusing of the flows and density increases in axisymmetric rings (V. Icke and G. Mellema, Leiden Observatory; A. Frank, Univ. Minnesota).

Circumstellar material lies in great quantity around stars of high masses. A star that forms at 60 solar masses may lose 55 solar masses and never become a cool extended star before it explodes as a supernova (N. Langer, MPIA, Munich). Such a star develops into a brief-lived 'luminous blue variable', of which only a half dozen are known in our Galaxy (L. Smith, University College London).

Luminous blue variables are related to Wolf-Rayet stars, which have massive wind outflows; but which are the mothers and which the children is not yet clear (Langer, Smith). Somewhat less-massive stars become blue or red supergiants, like the progenitors of SN1987A and SN1993J. A well-observed ring illuminated by SN1987A consists of matter ejected before the star exploded (C. Fransson and



Emission from the molecule MgNC ($N = 8-7$ transition) forms a stunningly detailed image of the source IRC+10216 made with the IRAM Plateau de Bure interferometer, whose 5-arcsec beam resolution is shown in the corner. The central star (marked with a cross), is a nearby carbon-rich, low-gravity, cool, giant star which is losing mass at a great rate, the material having been dredged up from within. The emission at the star and in a cloud around it comes from carbon-chain molecules such as C_2H , ..., C_6H , cyanopolyynes, $HC_{2n+1}N$, and metal-bearing molecules such as MgNC, SiS, SiC_2 and NaCl, composed principally of elements produced from hydrogen processed in nuclear reactions in the star's core. Emission from the cloud is layered in shells because of molecular abundance gradients and excitation conditions. (M. Guélin, IRAM, Grenoble.)

P. Lundqvist, Stockholm Observatory).

In a supernova explosion in a dense medium, the kinetic energy of the explosion is thermalized and radiated in a pulse as, for instance, with SN1987F; the spectrum, the pulse shape and the total energy involved resembled the spectra and flares in the nucleus of Seyfert galaxies (R. Terlevich, RGO, Cambridge; G. Tenorio-Tagle, IAC, Canary Islands). A pulse of heated discussion followed the proposition that some galactic nuclei are not powered by black holes but by numerous individual supernovae exploding into a dense medium — that is, by a large collection of fire-crackers instead of one superbomb. □

Harm Habing is at Leiden Observatory, Leiden, 2300RA, Netherlands. Paul Murdin is at the Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK.

DAEDALUS

Wet blanket

LAST week Daedalus remarked that sodium ions are mobile both in liquid water and in solids like beta-alumina. In water they are hydrated; in beta-alumina they are not. So in passing from the liquid to the solid, they would dehydrate, thus liberating water. Emerging back into the liquid, they would take up water again.

Daedalus now wants to drive this process with sunlight. He points out that solar photons have energies in the ion-dehydration range. At a solid-liquid electrolyte junction, they could dehydrate the water-borne ions and promote their entry into the solid. Sadly, the most promising ions and electrolytes for this process are colourless. To capture light for chemical purposes, the junction would need to be doped by a photosensitizing dye. So, says Daedalus, take a semipermeable tube of liquid electrolyte, join it at both ends to a rod of solid electrolyte, dope one junction with a suitable dye, and stand the result in sunlight. At the doped junction, ions will go from liquid to solid, dehydrating in the process. The released water will exude from the junction. At the undoped junction, the ionic current will emerge into the liquid again, rehydrating and taking up water. In the absence of any other source, it would even extract it from the air.

The applications are obvious. It is precisely the sunniest regions of the world that find it hardest to obtain water. But even the driest desert air has lots of water in it, which can now be simply captured. Daedalus hopes to miniaturize his photovoltaic vapour-condenser into a simple hollow fibre, whose outside absorbs water-vapour from the air, and pumps it into the central channel. Woven into fabric and exposed to sunlight, DREADCO's 'Sundew' fabric will deliver a steady trickle of captured pure water from its fibre ends. For the first time in the history of human settlements, water supply will no longer be a problem.

A simple Sundew-fabric tent will meet the watery needs of the most isolated explorer, desert wanderer or shanty-town dweller. As panels in a greenhouse, Sundew will capture water from the air outside, as well as condensing and recycling the transpired vapour from the plants inside. Sundew clothing will absorb evaporated sweat, and recycle it as potable water. Woven into sails, it will effortlessly solve the marine water problem. The nightmare technology of modern hydraulic engineering, the dams and reservoirs and vast webs of leaky pipes, as well as the endless wretched trudging for dirty well and river water throughout the third world, will be mercifully obsolete. David Jones

HIV-1 error revealed

SIR — We regret to inform *Nature's* readers that one experimental result in our recent Letter¹ is incorrect. The clones of one of the HIV-1 (human immunodeficiency virus type 1) mutants that we studied, called mutant 4, had unintended mutations in addition to the intended mutations in reverse transcriptase codons 74, 103, 215 and 219. We discovered these unintended mutations in reverse transcriptase only on further sequencing; they can explain why viruses derived from these clones were not viable. This explains the discrepancy between our report¹ and the findings of Larder *et al.*² and Emini *et al.* (see below). We believe that all other data in our report are correct, and we have confirmed the *in vitro* treatment results in subsequent experiments³.

We had hypothesized that one rationale for the use of several drugs directed against the same HIV-1 target was that some potential paths to multidrug resistance might not be available to the virus and never be seen under multidrug selection pressure. However, some mutations that confer resistance to non-nucleoside reverse transcriptase inhibitors can combine with AZT and ddI (2',3'-dideoxyinosine) resistance mutations to yield viable, triply resistant viruses *in vitro*. In extensions of the *in vitro* drug selection experiments we mentioned in our Letter⁴, we have also been able to select for triply resistant viruses similar to those described by Larder *et al.*².

Interactions between drug-resistance mutations will help to explain the structure and function of reverse transcriptase. However, caution is necessary in extrapolating from mutational interactions noted *in vitro* to the clinical potential of a particular combination. The primary criterion for choosing drugs to combine in clinical trials should be antiviral effectiveness. Our *in vitro* treatment experiments^{1,3} and clinical-trial data⁵ suggest that the combination of nucleoside and non-nucleoside reverse transcriptase inhibitors may provide prolonged antiviral effectiveness. Clinical trials comparing AZT, ddI and a non-nucleoside reverse transcriptase inhibitor to more standard combinations (such as AZT and ddI or AZT and ddC (2',3'-dideoxycytidine)) are now under way.

We agree with Emini *et al.* that multidrug-resistant variants are likely to be selected eventually during combination

therapy. But it remains important to explore new strategies for available drugs. Some drug combinations, particularly those which are more potent inhibitors of HIV-1 replication, may prolong the benefit of therapy more than others. This can only be determined by clinical trials.

Yung-Kang Chow

Martin S. Hirsch

Joan C. Kaplan

Richard T. D'Aquila

*Infectious Disease Unit,
Massachusetts General Hospital,
Boston,
Massachusetts 02114, USA*

HIV and multidrug resistance

SIR — The reverse transcriptase of human immunodeficiency virus type 1 (HIV-1) is inhibited by the nucleoside analogues AZT and 2', 3'-dideoxyinosine (ddI) as well as by the non-nucleoside compounds BI-RG-587 and L-697,661 (refs 1, 2). Unfortunately, the usefulness of these inhibitors as single therapeutic agents is limited by the emergence, in treated in-

series of mutant HIV-1 variants. The variants are listed in the table, together with their respective susceptibilities to the four test inhibitors. The combined effects of the expressed substitutions on virus susceptibility, particularly to AZT, are complex. For instance, we confirmed a previous observation that substitution at residue 74 resensitizes virus to AZT in the presence of the AZT resistance-engendering mutation at residue 215 (ref. 7). But the effect of the residue 74 alteration was abrogated in some cases (variants B and C but not E) by co-expression of the non-nucleoside inhibitor-specific substitutions at residues 103 or 106.

In any case, variants B and C were multiply resistant to AZT, ddI, L-697,661 and/or BI-RG-587, whereas variant D was resistant to ddI and both non-nucleoside inhibitors. Surprisingly, variant E is the identical putative nonviable variant constructed by Chow *et al.*⁶. In our hands, this mutant virus exhibited growth kinetics similar to wild-type virus (data not shown) and expressed a multiply resistant phenotype. We do not know why our observation differs from that of Chow *et al.* Nonetheless, we have demonstrated that multiply resistant viable HIV-1 variants can exist. The apparent phenotypic interactions that occur among co-expressed reverse transcriptase mutations reflect

SUSCEPTIBILITY OF HIV-1 VARIANTS TO TEST INHIBITORS

Variant	Amino-acid substitutions	IC ₉₅ (μM)			
		AZT	ddl	BI-RG-587	L-697,661
Wild type	None	0.025	12.5	0.400	0.100
A	74 (L→V), 215 (T→Y)	0.100	100.0	0.100	0.025
B	74 (L→V), 103 (K→N), 215 (T→Y)	50.0	100.0	>3.0	0.800
C	74 (L→V), 106 (V→A), 215 (T→Y)	50.0	100.0	>3.0	0.100
D	74 (L→V), 181 (Y→C), 215 (T→Y)	0.100	100.0	>3.0	>3.0
E	74 (L→V), 103 (K→N), 215 (T→Y), 219 (K→Q)	0.100	100.0	>3.0	0.800

Variant proviruses were constructed and 95 per cent virus inhibitory concentrations were determined as described in ref. 3.

fectured persons, of resistant virus variants (see refs 3–5). Chow *et al.*⁶ have proposed the concept of “convergent combination therapy”, that simultaneous multiple inhibitor therapy directed against the HIV-1 reverse transcriptase would delay the emergence of multiply resistant virus strains.

Given the contrived nature of the cell-culture experiments of Chow *et al.*⁶ designed to select resistant virus, we decided to determine directly if multiply resistant virus could indeed exist by constructing a

significant structural and functional flexibility by the enzyme. This flexibility permits the expression of virus variants that can be resistant to various combinations of reverse transcriptase inhibitors. Given the large quantities of virus seemingly present in infected individuals⁸, such variants are likely to be selected during multiple therapy.

Emilio A. Emini, Donald J. Graham,

Leah Gotlib, Jon H. Condra,

Vera W. Byrnes, William A. Schleif

*Merck Research Laboratories,
West Point, Pennsylvania 19486, USA*

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references.

1. Chow, Y.-K. *et al.* *Nature* **361**, 650–654 (1993).
2. Larder, B. A., Kellam, P. & Kemp, S. D. *Nature* (in the press).
3. Mazzulli, T., Chow, Y.-K., Merrill, D. P. & Hirsch, M. S. Abstract PO-B26-2082 IX Int. AIDS Conf., Berlin (1993).
4. Chow, Y.-K., Merrill, D. P., Kaplan, J. C., D'Aquila, R. T. & Hirsch, M. S. Abstract WS-A19-6 IX Int. AIDS Conf., Berlin (1993).
5. Havir, D. *et al.* Abstract WS-B26-1 IX Int. AIDS Conf., Berlin (1993).

1. Goldman, M. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 6863–6867 (1991).
2. Merluzzi, V. J. *et al.* *Science* **250**, 1411–1413 (1990).
3. Byrnes, V. W. *et al.* *Antimicrob. Ag. Chemother.* (in the press).
4. Richman, D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 11241–11245 (1991).
5. Nunberg, J. H. *et al.* *J. Virol.* **65**, 4887–4892 (1991).
6. Chow, Y.-K. *et al.* *Nature* **361**, 650–654 (1993).
7. St Clair, M. H. *et al.* *Science* **253**, 1557–1559 (1991).
8. Piatak, M. *et al.* *Science* **259**, 1749–1754 (1993).

Ageing without sex?

SIR — Is it possible to grow old without having sex? This intriguing question was discussed in a stimulating review¹ in which Partridge and Barton argue that accumulation of deleterious mutations is at least a part of the cause of senescence. In asexual lineages deleterious mutations inevitably accumulate by 'Muller's ratchet'². This happens because, in a finite population, there is always a chance that the class of individuals with the fewest deleterious mutations will be lost. Without recombination, this class cannot be recovered and so fitness (or longevity in this case) will decline irreversibly.

Discussion of Muller's ratchet always raises the question of persistent asexual lineages²; if they exist in organisms with large genomes and population sizes less than about 10^{10} (ref. 3) they present a paradox. Partridge and Barton refer to lineages of fish and salamanders, believed to be aged 4–5 million and 100,000 years old, respectively, based on mitochondrial DNA analyses, but argue that they are not truly asexual. A more commonly quoted example is the bdelloid rotifers, a whole order of animals that is considered to be entirely asexual². However, the rotifers have essentially no fossil record⁴ and have yet to be subjected to molecular analysis.

We wish to draw attention to the ostracod superfamily Darwinuloidea. These freshwater crustacea have an excellent fossil record extending back to the Carboniferous (>286 million years). They are very distinct in that females possess a brood pouch, absent in most other freshwater ostracods; not surprisingly, pouch-brooding leads to marked shell sexual dimorphism in other ostracod groups (such as Timiriasevinae). Modern darwinuloids are exclusively parthenogenetic, and there is no evidence of shell dimorphism (or sexuality) in any fossil darwinuloid since the late Mesozoic (about 70–90 million years ago)⁵. There are two extant genera, *Darwinula* and

Microdarwinula, in which all species are exclusively asexual. *Darwinula* spp. are known into the Carboniferous⁶, and one species, *D. stvensoni*, is common in deposits throughout the Quaternary and earlier into the Miocene (24.6 million years ago), but a single report of a male now appears unreliable⁶.

D. stvensoni lives for 3–4 years, of which it may be adult for 2 years^{7,8}. This is considerably longer than most freshwater ostracods, whose life cycles typically last for 1 year at most⁹. It is not rare or obscure as population densities may exceed 10^5 m⁻² (ref. 8). Therefore it does appear to represent a paradox.

Large population size may be a way out of this paradox. Populations of *D. stvensoni* may exceed the limit of 10^{10}

above which Bell³ considers sex unnecessary to avoid Muller's ratchet, but these large population sizes are unlikely to be maintained in the shallow freshwater habitat of the species for extended periods of time. Certainly, populations present in Europe today must have passed through much smaller population sizes during periods of postglacial expansion.

Maynard Smith says we 'badly need to know' how bdelloid rotifers manage without sex². Given their excellent fossil record, perhaps attention should also be paid to the Darwinuloidea.

R. K. Butlin

H. I. Griffiths

Department of Genetics,
The University of Leeds,
Leeds LS2 9JT, UK

Ancient use of cannabis

SIR — From the presence of *Cannabis sativa* in the ancient tomb of a girl who probably died in childbirth, Zias *et al.* in their Scientific Correspondence¹ conclude that the plant was used to increase the force of uterine contractions and decrease labour pains, and was administered by a midwife because "in antiquity" physicians were prohibited by law from attending women in labour. Although we agree that the plant probably was used for medicinal purposes, we do not think that it was used for the purpose Zias *et al.* suppose.

Even if we assume that cannabis does increase the force of uterine contractions and decrease labour pains (we are not aware of recent studies on these subjects), the fact that it was used in a difficult delivery does not prove that it was used because of these pharmacological properties. In ancient medicine, many plants were used (of the more than 300 medications mentioned in the *Corpus Hippocraticum* about 250 are plants²), and therefore it would not be surprising if cannabis were found among the plants used in the case discussed. But this does not mean that it was used because of its (alleged) effects on the uterus.

As a general rule, if a certain plant used in antiquity has a particular pharmacological property, it is not justified to assume that it was used because of that property³ unless it can be shown that the drug was used more often than others in the appropriate clinical situations⁴. If, for example, besides Δ^9 -THC, Zias *et al.* had also found traces of opium, would they be justified in concluding that the patient was given *Papaverum somniferum* to relieve her pain? No. Taking the *Corpus Hippocraticum* as a guide, it is clear that the plant was often used, but usually with other plants and not for the treatment of pain. Rather it was used for such disparate conditions as empyema, "phthisis", "typhus", leucorrhoea, habitual abortion,

"dropsy of the uterus", metrorrhagia and "displacement of the uterus". In all these cases there is no mention of pain. On only three occasions is the relief of pain mentioned and, in these, poppy is used as one ingredient in a mixture with other drugs (thirty in one case, four — burned for fumigation — in another) and only in one case is given with just one other drug. Therefore, we cannot conclude that poppy was used by Hippocratic physicians for the purpose of controlling pain simply because they used it and because we know that it contains opium. Similarly, as we do not know the circumstances concerning the use of cannabis in the case described by Zias *et al.*, it is likely that it was used (probably with other herbs) by chance and not for its (alleged) effects on labour pains.

We do not know why Zias *et al.* believe that "in antiquity" physicians were prohibited by law from attending women in labour, but in ancient Mesopotamia the physician presumably attended deliveries as he performed caesarean sections⁵. In the *Corpus Hippocraticum* there is ample evidence that physicians attended deliveries, and Soranus' *Gynecology* dispels any notion that physicians did not attend women in labour in the Roman world.

Plinio Prioreschi

Division of History of Medicine,
Department of Pharmacology,

Donald Babin

Department of Biochemistry,
Creighton University School of Medicine,
Omaha,
Nebraska 68178, USA

1. Zias, J. *et al.* *Nature* **363**, 215 (1993).
2. Von Staden, H. *Herophilus: the Art of Medicine in Early Alexandria* (Cambridge University Press, 1989).
3. Prioreschi, P. *A History of Medicine, Volume I, Primitive and Ancient Medicine* (Mellen, 1991).
4. Estes, J. W. *The Medical Skills of Ancient Egypt* (Science History, 1989).
5. Oppenheim, A. L. *J. Hist. Med. Allied Sci.* **XV**, 292–294 (1960).

1. Partridge, L. & Barton, N. H. *Nature* **362**, 305–311 (1993).
2. Maynard Smith, J. in *The Evolution of Sex* (eds Michod, R. E. & Levin, B. R.) 106–125 (Sinauer, Sunderland, Massachusetts, 1988).
3. Bell, G. in *The Evolution of Sex* (eds Michod, R. E. & Levin, B. R.) 126–138 (Sinauer, Sunderland, Massachusetts, 1988).
4. Briggs, D. E. G. & Crowther, P. R. *Palaeobiology: a Synthesis* (Blackwell, Oxford, 1990).
5. Whitley, R. in *Aspects of Nonmarine Cretaceous Geology* (eds Mather, M. J. & Chen, P.-J.) 177–192 (China Ocean, Beijing, 1992).
6. Sohn, I. G. *Micropaleontology* **33**, 150–163 (1987).
7. McGregor, D. L. in *The Taxonomy, Morphology and Ecology of Recent Ostracoda* (ed. Neale, J. W.) 194–221 (Oliver & Boyd, Edinburgh, 1969).
8. Ranta, E. *Ann. Zool. Fenn.* **16**, 28–35 (1977).
9. McLay, C. L. *Can. J. Zool.* **56**, 1170–1179 (1978).

The measure of the man

William R. Shea

Newton for Beginners. By William Rankin. *Icon Books, 45 Princess Road, London NW1 8JS, UK: 1993. Pp. 176. £7.99 (pbk).*

The Life of Isaac Newton. By Richard S. Westfall. *Cambridge University Press: 1993. Pp. 328. £16.95, \$24.95.*

Equivalence and Priority: Newton Versus Leibniz. By Domenico Bertoloni Meli. *Oxford University Press: 1993. Pp. 318. £55, \$72.*

Fits, Passions and Paroxysms. By Alan E. Shapiro. *Cambridge University Press: 1993. Pp. 400. £45, \$69.95.*

All Was Light: An Introduction to Newton's Opticks. By A. Rupert Hall. *Oxford University Press: 1993. Pp. 252. £35, \$58.*

THE newtonian industry is thriving and there is no recession in sight. Younger scholars, such as Domenico Bertoloni Meli and Alan E. Shapiro, have made capital investments of time and research that are beginning to pay handsome dividends. Among more senior scholars, Richard S. Westfall is finding a broader market for his earlier work by reissuing it in a more accessible form, while A. Rupert Hall offers a new product whose content and packaging are outstanding.

Shoddy imitations of the genuine article are also beginning to appear, as readers of the cartoon guide *Newton for Beginners* will discover to their annoyance. The author's obvious dread of being considered a bore is not entirely misguided. He will do anything to get a laugh. For instance, when the proverbial apple lands on his head, Newton is not made to utter a word about gravity but says: "That reminds me. I must feed the cat."

The real book for beginners is Westfall's *The Life of Isaac Newton*, a summary of his monumental biography *Never at Rest*, which was first published in 1980 and ran to 908 pages. Westfall is a lively and lucid expositor of Newton's ideas and he has written an excellent book for those who want an authoritative introduction to Newton but do not have the time or the inclination to wrestle with the finer points of his mathematics. There is indeed more to Newton than the method of fluxions or the law of universal gravitation and, with the exception of the 18-month period (1665–66) when he laid the foundation of his calculus and his cosmology, and the two-and-a-half years (1684–86) when he wrote the *Principia Mathematica*, Newton spent much of his time on subjects that required little or no mathematics. He performed a wide variety of chemical and optical experiments, tracked down the path of light rays, ground lenses and mirrors to build telescopes and, above all, read voraciously in the realms of theology and alchemy. Newton shared the general assumption of his age and expected ancient texts to provide not only the basis of a literary and moral formation but scientific and technical instruction as well.

Westfall is particularly illuminating about Newton's interest in the chemical

raciously in the realms of theology and alchemy. Newton shared the general assumption of his age and expected ancient texts to provide not only the basis of a literary and moral formation but scientific and technical instruction as well.

Westfall is particularly illuminating about Newton's interest in the chemical

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Sir Isaac Newton (1642–1727).

philosophers for whom the unity and harmony of the Universe was an obvious consequence of creation by a wise and benevolent deity. The records of Newton's alchemical experiments are not always couched in the imagery common to alchemical writers, but Westfall argues that they were performed against the background of two of their tenets: first, the belief that generation by male and female is a universal process at all levels of nature, and, second, the need to purge and cleanse in order that the spiritual seeds of things may reach maturity. I should mention that *The Life of Isaac Newton* has no footnotes, and readers who wish to find the source of a particular

quotation will have to look it up in *Never at Rest*.

Meli's *Equivalence and Priority: Newton Versus Leibniz* was born as a doctoral dissertation and retains the characteristic strengths of the genre: the fastidious footnotes display a mastery of the entire bibliography of seventeenth-century mechanics, and passages from Leibniz's working notes are subjected to a precise, even pitiless scrutiny. The priority dispute does not refer to the discovery of calculus but to the inverse square law, familiar to later generations as $m_1 m_2 / r^2$. The core of Meli's book is an examination of Leibniz's reaction to the publication of Newton's *Principia* in 1687, as revealed in three short essays that appeared in the *Acta Eruditorum* in 1689 and three brief posthumous texts. The third of these texts consists of notes that Leibniz scribbled when reading the *Principia* and they are reproduced here for the first time in a painstaking critical edition.

Newton, it will be remembered, postulated that the planets move around the Sun under the action of gravitational attraction whereas Leibniz, following Descartes' suggestion, assumed that they were moved by celestial matter on the analogy of the revolution of pieces of straw in a whirlpool. In one of the three essays he published in 1689, Leibniz expressed himself as though he had come across only a review of Newton's *Principia* and had not yet seen the book. It would seem therefore that Leibniz's manuscript notes of the *Principia* were written after he returned from an extended tour of libraries in southern Germany, Austria and Italy between November 1687 and June 1690. Leibniz's manuscript consists of only two sheets folded in quarto. The first has no watermark, the second bears one that is identical to the kind found on paper that Leibniz used in Vienna in 1688. On this slender evidence, Meli concludes that Leibniz saw

the *Principia* in that city in that year. Had Meli been writing from Brussels rather than Cambridge, he might have been reminded that people who travel carry paper along with them and that Leibniz could easily have used in 1690 a sheet from notepaper acquired a couple of years earlier. (It so happens that the sheets on which I am typing this review were purchased four years ago.)

The *Fits, Passions and Paroxysms* of Shapiro's book do not refer to Newton's emotional state but to his description of the properties of light in thin films and thick plates. Newton suggested in his *Opticks* that light rays are put into periodically recurring states (fits) that dispose

them to be easily reflected or transmitted. From the time it was first proposed, the suggestion had few passionate supporters and has often been described as obscure or strange. As the editor of Newton's *Optical Papers* and an authority on the history of light, Shapiro is well prepared to write a monograph on Newton's theory. Indeed, he does more, and we find not one but two monographs between the covers of this book. The first is a study of Newton's theory of coloured bodies, the second a survey of its reception, challenge and eventual transformation between 1704 and 1833.

Those who want a more general introduction to Newton's *Opticks* can turn to Hall's *All Was Light*, which captures both the intellectual daring and the experimental genius of Newton. Hall offers a detailed reconstruction of the development of Newton's investigation of light from the seminal period around 1666 when he discovered that sunlight consists of a mixture of rays of different colours that differ in refrangibility, the red ones less refracted than the violet at the other end of the spectrum. Soon afterwards, Newton discovered a method for measuring the thickness at which colours appear in a thin film. This discovery of what were afterwards called 'Newton's rings' was

communicated to the Royal Society in 1675 but not set out fully until the publication of his *Opticks* in 1704.

The care and patience with which Newton carried out his experiments give us the measure of the man. He placed a lens of known curvature on a flat piece of glass and observed the coloured rings that appeared when he looked through the lens. Using an ordinary compass and the naked eye, he recorded one circle at $23\frac{1}{2}$ hundredths of an inch in diameter and the next at $34\frac{1}{3}$. When a small discrepancy appeared in his results, he relentlessly stalked the source until he found that the two faces of his lens had slightly different curvatures. The difference corresponded to a measurement of less than one-hundredth of an inch in the diameter of the inner circle and about two one-hundredths in the diameter of the sixth. Such were Newton's exacting standards of accuracy. □

William R. Shea is at the Centre for Medicine, Ethics and Law at McGill University, Montreal, Quebec, Canada H3A 1W9.

■ A. Rupert Hall has also recently written a biography of Newton, *Isaac Newton, Adventurer in Thought* (Blackwell, 1992). It was reviewed in *Nature* 363, 29 (1993).

Manufacturing knowledge

Stuart Blume

Making Science: Between Nature and Society. By Stephen Cole. *Harvard University Press*: 1992. Pp. 290. \$39.95, £31.95.

WHEN the sociology of science emerged in the 1960s as an area of professional sociological inquiry, it did so in sharp distinction to the existing discipline of history and philosophy of science. The content of science, scientific knowledge, was seen as principally a consequence of the organization of the natural world and hardly subject to sociological investigation. Questions of the genesis and status of facts and theories were matters for historians and philosophers. Inspired principally by the dominating figure of Robert K. Merton (professor of sociology at Columbia University), sociologists set out to investigate clearly social aspects of the behaviour of scientists and the functioning of scientific institutions. Merton had argued long before that certain 'norms' governed the way in which scientists collectively behaved. Many of his students addressed their research to the empirical consequences of Merton's ideas: for example, how far did scientists in practice evaluate one another's work purely in terms of its inherent scientific worth? One of these students, now himself a professor of sociology, was Cole.

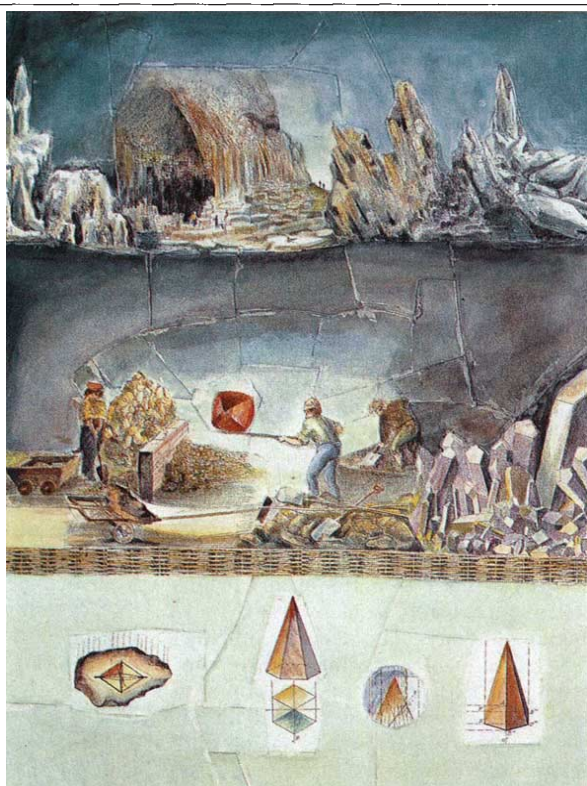
A good example of the work done in this tradition was a study of the working of the peer review system at the National Science Foundation which Cole and others carried out in the mid-1970s. Reviewing 'phase 1' of this study in *Nature* 287, 807 (1980) I wrote: "An effective combination of statistical and qualitative analysis is used to show, most simply, that the best research gets funded. No evidence for an old boy network is uncovered, and the findings suggest that referees are not much influenced by the status of the applicant's institutional affiliation, or by his seniority. They simply judge the application on its merits". But by the time this study (and my review) was published, sociologists, influenced by the work of Kuhn, were starting to ask quite different questions about science.

Sociologists began to articulate a sociology of scientific knowledge, in which scientific theories and methods were no longer to be viewed as uniquely related to the nature of the natural world. The statements articulated and advanced by scientists began to be viewed as products of the social processes through which articulation and so on took place. A new sociology of science was emerging,

Crystal scale

Giving in

At 1.4 million atmospheres xenon, a gas, goes metallic. Between squeezed single-bevel diamond anvils, jagged bits of graphite shot with a YAG laser form spherules. No one has seen liquid carbon. Try to imagine that dense world between ungiving diamonds as the pressure mounts, and the latticework of a salt gives, nucleating at defects a shift to a tighter order. Try to see graphite boil. Try to imagine a hand, in a press, in a cellar in Buenos Aires, a low-tech press, easily turned with one hand, easily cracking a finger in another man's hand, the jagged bone coming through, to be crushed again. No. Go back, up, up like the deep diver with a severed line up, quickly, to the orderly world of ruby and hydrogen coloring near metallization, but you hear the scream in the cellar, don't you, and the diver rises too fast.



The amalgam of science and art displayed here is taken from *Chemistry Imagined*, an unusual melding of short essays and poems by the Nobel laureate Roald Hoffmann with 30 colour collages by Vivian Torrence. The aim of the volume, which arose from an exhibition of the same title, is to convey the creative spirit of chemistry through the ages. Smithsonian Institution Press, \$19.95, £14.95.

exemplified initially in the work of S. B. Barnes, D. Bloor and M. Mulkay and in books such as B. Latour and S. Woolgar's *Laboratory Life* (Sage, 1979/Princeton, 1986). In this new (largely European) sociology of science, in which scientific knowledge was seen as 'socially constructed', the questions and the methods that Merton's students had pioneered faded from view.

Cole's new book is an attempt at a resurrection. He wants to show the limitations of current constructivist sociology and the need to reintroduce the older perspective. A large part of the book consists of a re-presentation of his own and his colleagues' earlier work (such as the study of peer review), but now as though it were addressing the current scholarly agenda. It is an oddly painful intellectual exercise. "All constructivists", we are told, "argue that the actual cognitive content of the natural sciences can only be understood as an outcome of social processes and as influenced by social variables." The constructivists' 'claim' can then be 'tested' by re-examining data on the extent to which 'social variables' (such as working at a leading university) influence the reception of a piece of scientific work. Few sociologists, still fewer historians of science, will accept the reduction of the complex patterns of behaviour that figure in their own accounts to the standard variables (age, affiliation and so on) that figure in Cole's. The logic of this attempted resurrection rests on implausible interpretations, readings and distinctions.

It may be, Cole grants, that what scientists do in their laboratories is not precisely determined by nature, that it involves all kinds of negotiations and other social processes. But surely, what finds ultimate acceptance within the scientific community is not a matter only of rhetoric or power? Surely nature plays some part; so surely the constructivists are just wrong — the 'constructivists' with whom Cole purports to be debating are a feeble compound, a parody.

There are many whose researches today are directed towards understanding how social processes are implicated in the detail of scientific work and who reject totally the idea that it all boils down to power, rhetoric or negotiation. There are many sociologists of science who believe that their field is ripe for another shift, and that this might well entail looking anew at the institutions, the responsibilities and the careers of scientists. Unfortunately, Cole's attempt at resurrection, which is unconvincing and unfortunate, has done intellectual renewal no service. □

Stuart Blume is in the Department of Science Dynamics, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands.

Earth's twin

David W. Hughes

The Evening Star: Venus Observed. By Henry S. F. Cooper Jr. Farrar, Straus & Giroux: 1993. Pp. 273. \$22.

EARTH is 5.4 per cent bigger, 5.1 per cent denser and 23 per cent more massive than Venus. But what are a few per cent to an astronomer? In astronomical circles the planets are nearly as close as they can get to being twins. It is only their distances from the Sun that differentiates them. Venus is closer, and therefore hotter. These things have got out of hand. A 91 per cent increase in the intensity of incident solar radiation has seemingly given Venus a surface temperature of 460 °C and has led to the production of an atmosphere of carbon dioxide that exerts a pressure some 90 times greater than the pressure at the surface of Earth, equivalent to being 2,700 feet deep in the ocean.

If astronauts could survive being both squashed and boiled they would, when standing on Venus, see a dark reddish-brown rocky landscape languishing under a uniformly luminous overcast sky that was occasionally enlivened by a flash of lightning. Only a breath of torrid wind would stir the air.

The big questions concern the evolution of the surface. Has the high temperature and pressure, and the absence of water, given Venus a completely different surface from that of Earth? Has the crust of Venus split into plates? Do these roam about and collide, producing mountain chains and volcanism? Does the venusian volcanism come from another source? Is there any granite or are all the rocks basalt? Is the average age of the surface 500 million years? Has only 10 per cent volcanically resurfaced since then?

Planetary twins seem to be more fascinating than our more disparate siblings and this has led to Venus being visited by more spacecraft than any other planet. Henry Cooper writes about the Magellan mission. This shuttle-launched NASA spacecraft left Earth in May 1989 and on 10 August of that year started to orbit Venus 7.3 times a day, dipping at its lowest to a height of 265 km above the planet's surface. The on-board synthetic aperture radar could resolve surface features to an accuracy of 120 metres. The altimeter added the third dimension. The second cycle of observations produced stereoscopic images.

The Magellan mission has been so suc-

cessful that it will surely spawn many books, but I am convinced that there will be none quite like Cooper's. The Magellan mission has resulted in sheaves of surface images that have been pored over by hundreds of geologists and geophysicists. Pictures are everything to these scientists, but apart from two end-page maps, this book is unillustrated. Time and time again the text gives the impression of a Frenchman trying to describe a spiral staircase with his legs tied together and his hands glued into his pockets.

Cooper has done his best to describe the scientific arguments that raged during the data analysis at the Jet Propulsion Laboratory but, with his qualifications in English

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

False-colour image of Venus's Eistla Regio volcanic edifice, made by the Magellan spacecraft in March 1991. The slightly concave summit is 35 km in diameter. Black area has insufficient data.

literature and background on *The New Yorker*, he is reluctant to be partisan or to push forward any of his own ideas. One often comes across sections such as "XXXX agreed with this assessment. I asked him whether YYYY's rigid lithosphere was consistent with the observed plasticity of the crust" — followed by a page or so of quoted response.

The book only really leaps to life when the author discusses scientists rather than science. The anger and the arguments, the petty feuds, the politics of science, the grubbing for grants, and the trifling jealousies make fascinating reading. Cooper has provided some 270 pages of 'fly on the wall' journalism applied to the space-science data-handling business. □

David W. Hughes is in the Department of Physics, University of Sheffield, Sheffield S3 7RH, UK.

Principia Microscopica

Graham Dunn

Light and Electron Microscopy. By Elizabeth M. Slayter and Henry S. Slayter. Cambridge University Press: 1993. Pp. 256. £35, \$64.95 (hbk); £14.95, \$27.95 (pbk).

WHAT a refreshing change to open a new book on microscopy that is not only packed with information on recent advances in the field but also contains a broad foundation of theoretical principles. The volume is slim, and it is hard to believe that the wide range of material is covered in such depth until one realizes what has been left out: the usual dreary rules-of-thumb for operating microscopes that might as well be describing masonic rituals for all the explanation — or intellectual stimulation — that they offer. In fact, I opened the book with pleasurable anticipation because Elizabeth Slayter was already familiar to me as the author of my well-thumbed copy of *Optical Methods in Biology* (Wiley, 1970). Her untimely death in 1987 is a sad loss to light microscopy. Her coauthor here, also her husband, attributes "any clarity, thoroughness, or felicitous style" to her; but the consistent expression of these qualities suggests he is too modest.

Although the book is not targeted exclusively at biologists, it is assumed that it "will appeal especially to graduate students and researchers in the biomedical sciences". But because most biologists are not proficient in the theory of vector fields, it may seem risky to kick off with anything as intimidating as Maxwell's equations. Is this the fabled *pons asinorum* that confronts us or are we being invited simply to marvel at the cornerstone of modern optical theory? Probably the latter, because whenever mathematical notation is introduced elsewhere, the concepts are explained verbally in the best tradition of Richard Feynman's *Lectures on Physics*. Most research biologists should be able to cope with the text but there are undoubtedly some bridges that mathematical asses will find difficult to cross. Those with an incurable terror of equations should resign themselves to the fact that microscopy, in theory if not in practice, is essentially a physico-mathematical subject.

Having negotiated Maxwell's equations by whatever means at his or her disposal, the reader is coaxed through the principles of microscope optics and the Fourier theory of image formation. I envy those who first encounter these topics here, for the clouds of incomprehension are swept away so deftly that enlightenment may

creep up on them unawares. It will probably begin to dawn half way through the book, in a chapter disarmingly called "The Light Microscope", when the familiar features of design and operation of the microscope will suddenly be seen in a new light — the light of understanding.

The remaining chapters largely describe how the principles are applied in a wide variety of microscope designs. Not only are the usual forms of light and electron microscopy covered comprehensively, with special emphasis on biological applications, but novel instruments such as the total-internal-reflection microscope are also described briefly. Indeed, the book explores regions well beyond the bounds set by its title, devoting a whole chapter to X-ray microscopy. Another chapter presents a veritable menagerie of such exotic species as field-ion, neutron, photoelectron, scanning tunnelling, atomic force, nuclear magnetic resonance and acoustic microscopes. The final chap-

ter concerns photography, seemingly mundane after all this high technology, but even this familiar topic is treated with a deep insight into underlying principles.

Microscopy is currently enjoying a renaissance as new approaches to image formation, advances in digital image processing and developments in nanotechnology take it into areas of application hitherto undreamed of. Even routine forms of microscopy are now changing rapidly and there is greater need than ever for research biologists to understand the function of their instruments, if only to interpret the image that they produce. Although there is no easy route to understanding, this book will make the effort rewarding, and I suspect that it will soon find its way onto the shelves of microscopists who take their craft seriously. □

Graham Dunn is in the MRC Muscle and Cell Motility Unit, King's College London, 26–29 Drury Lane, London WC2B 5RL, UK.

Responses to change

J. John Lowe

Quaternary Environments. By M. A. J. Williams, D. L. Dunkerley, P. de Deckker, A. P. Kershaw and T. Stokes. Edward Arnold: 1993. Pp. 329. £16.95 (pbk).

SEVERAL textbooks on Quaternary geology or environmental history have appeared in recent years, and the authors of each have faced the same dilemma: how to get across the breadth and scope of this vast (and still growing) subject within the confines of a standard text. The answer, of course, is to avoid any attempt to 'cover' the subject, and instead to concentrate on key concepts, illustrate the astonishing rapidity and complexity of environmental changes during the Quaternary period, and exemplify the multidisciplinary nature of current research in the field. Authors of previous texts have attempted this with varying degrees of success. This new publication, written by teachers and researchers of the Australian National University, Monash University and the University of Adelaide, is one of the more successful efforts.

As an introductory text for undergraduate students, it has a very wide brief. It aims to examine global environmental fluctuations of the past 2 million years, establish the basis of some of the important methods of analysis used to reconstruct them, and consider the ways in which organisms (including humans) have responded to the fluctuations. It is inevitable, then, that much of the material lacks detailed explanation and that some important issues are oversimplified. But what the volume does do is to provide students

with a well-structured platform for understanding the pivotal role of Quaternary studies in science today.

The importance of understanding oceanic circulation, rapid environmental changes in today's desert regions, the history of Quaternary glaciations, the kaleidoscopic patterns of adjustment in global flora and fauna, theories of human origins and migrations, atmospheric circulation processes, and a knowledge of Quaternary environmental history as a necessary framework for assessing impacts of future climatic change — all of these are neatly introduced without a heavy terminological armour.

Particular attractions of the book are the treatments of topics that are dealt with too briefly or not at all in other texts. Thus the 'prelude' provided of global changes during the Tertiary, the detailed examination of evidence from desert regions and lakes in low latitudes, the useful models of oceanic and atmospheric processes and the succinct introduction to human history are all very welcome. The Southern Hemisphere perspective is also a bonus, and serves to remind northern readers that the North Atlantic is only a cog — albeit a large one — in an intricate global machine.

Overall, this is a lucid and lively book with compelling messages. It illustrates why the Quaternary period offers such a popular and vital forum for Earth scientists today. □

J. John Lowe is in the Centre for Quaternary Research, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.

Structure of gelsolin segment 1-actin complex and the mechanism of filament severing

P. J. McLaughlin, J. T. Gooch, H.-G. Mannherz* & A. G. Weeds

MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

* Institute of Cytopathology and Cytopathology, Philipps-University, Marburg, Germany

The structure of the segment 1 domain of gelsolin, a protein that fragments actin filaments in cells, is reported in complex with actin. Segment 1 binds monomer using an apolar patch rimmed by hydrogen bonds in a cleft between actin domains. On the actin filament model it binds tangentially, disrupting only those contacts between adjacent subunits in one helical strand. The segment 1 fold is general for all segments of the gelsolin family because the conserved residues form the core of the structure. It also provides a basis for understanding the origin of an amyloidosis caused by a gelsolin variant.

In most eukaryotic cell types, actin filaments are non-covalently crosslinked in isotropic-gel domains that constitute a clear layer between cytoplasm and cell-membrane. As a gel resists deformation, this actin cortex plays an important role in determining cell shape. Thus local fragmentation of actin filaments (F-actin) allows remodelling of the cortical layer to form the fluctuating structures seen at cell surfaces, such as protrusions or invaginations of cytoplasm. In cells, actin structures are indeed dynamic¹ and are affected by extra cellular signals². Gelsolins, and related proteins, are prime candidates for effecting such regulation.

Gelsolins have been implicated in actin-filament reorganization during cell locomotion³ or in vesicle traffic through the cortex⁴. A secreted form, with an additional 21-residue leader sequence⁵, is found in blood plasma. On Ca^{2+} activation, gelsolin converts actin gels to a sol by randomly fragmenting filaments⁶ (Fig. 1a). After breaking a filament, gelsolin remains tightly bound to one of the newly created ends, trapping a Ca^{2+} ion in the complex⁷, and acting as a 'cap' that prevents reannealing. The filament end blocked by gelsolin is the preferred end to which fresh monomers bind as new filaments extend from the fragmented nuclei. This gelsolin cap can be dissociated, *in vitro*, by vesicles containing the lipid phosphoinositide PIP_2 (ref. 8; Fig. 1a). As PIP_2 levels in the plasma membrane are altered by activation of receptors, such a mechanism would localize the growth of new filament structures close to the site of action of extracellular signals².

Gelsolin is a member of a family defined by a repeated sequence motif⁹ of 125–150 residues. The sequence motif is found in proteins from widely separated phyla (including amoebae, insects, birds and mammals) as either three repeats, or six (like gelsolin). Each repeat or segment probably represents an independently folded domain, because limited proteolysis occurs only at segment boundaries^{8,10,11} and the products obtained retain distinct actin binding properties. Like gelsolin, segment 1 binds monomeric actin very tightly and also traps a Ca^{2+} ion in the complex¹². The primary role of this segment is shown by the reconstitution of severing activity in an engineered chimera¹³ in which segment 1 is spliced onto the unrelated F-actin binding domain of α -actinin. A model for severing was proposed in which segment 1 binds tightly to an actin subunit and disrupts actin-actin contacts.

This specific case of actin severing has wider relevance: how are associations broken in a cooperatively assembled polymer? A model for F-actin has been derived from fibre diffraction data¹⁴ which provides a structural context in which to understand segment 1 binding to subunits. Each actin subunit makes

contact with four others in the double-helical filament. Does severing require disruption of only one subunit:subunit contact or might a severing agent bind at a critical point where more than two subunits abut? In addition to the gelsolin family, over 100 different actin-binding proteins are known: sequence homologies have been found between gelsolins and some of these other classes which may denote common modes of binding^{15,16}.

These questions can now be addressed as we have solved the structure of gelsolin segment 1 in complex¹⁷ with actin. In addition, analysis of the segment fold provides a molecular explanation for the aetiology of a familial amyloidosis (Finnish type), a hereditary disease caused by a point mutation in gelsolin.

Structure determination of the segment 1 fold

Crystals of human gelsolin segment 1: actin complex diffract weakly to 2.5 Å; a combination of short-wavelength synchrotron radiation and a large-area, image-plate detector was essential to measure adequate diffraction data (Table 1). The structure was solved by first locating actin in the unit cell by molecular replacement. After solvent flattening of the initial map, side chains could be placed for 42 residues in the structural core of segment 1. A model for segment 1 has been completed by subsequent rounds of manual building and refinement; significant changes in the actin structure illustrate independence of the refined phase set from the starting model. The current *R*-factor is 0.187 using all data to 2.5 Å (with a bulk solvent correction) for 476 residues, and 52 ordered water molecules.

Segment 1 is a globular domain with rough dimensions 13 Å × 26 Å × 25 Å. It is organized as a three-layered structure (Fig. 1b); the central four-stranded β -sheet is sandwiched between a four-turn α -helix that runs roughly parallel to the strands in the sheet and, on the other face, a shorter α -helix that runs roughly perpendicular to the strands. Conserved residues present in all segments of the gelsolin family contribute to the apolar core of the molecule¹⁸ (Fig. 1c). Most of these are in the middle two strands of the sheet and the shorter α -helix that packs on top of them. Two conserved carboxylate residues are also completely buried. None of the residues in invariant positions directly contacts actin, despite the fact that even surface residues in actin sequences are highly conserved across diverse phyla¹⁹. Rather, residues conserved in gelsolin family segments appear to reflect selective pressure to maintain the close-packed apolar core that defines the three-dimensional fold.

Actin binding site

The longer α -helix in segment 1 binds in a cleft formed at the

interface of actin subdomains 1 and 3 (Fig. 2a). A central patch of exclusively apolar atoms is rimmed by polar, hydrogen-bonding atoms, rather like a circular sticking plaster, where a gauze pad (apolar patch) is bordered by an adhesive edge (hydrogen-bonding residues) (Fig. 2b). On segment 1 the apolar patch consists of residues discontinuous in the linear sequence and centred

on helix residue Ile 103, which protrudes furthest from the surface. This branched apolar side chain inserts between apolar residues from both actin subdomains 1 and 3 (Fig. 2c). Hydrophobic residues from actin subdomain 1 are on the solvent-exposed face of a helix (341'–349') previously noted as unusually apolar in the actin:DNase 1 structure²⁰. Their conformations

TABLE 1 Data collection and refinement statistics

Crystal parameters:						
Abbreviation						
natSEG1:ACTIN						
Segment 1:actin complex						
Spacegroup $P2_12_12_1$ $a=57.4$ Å, $b=70.4$ Å, $c=185.4$ Å						
cysSEG1:ACTIN						
(C57N Segment 1):actin complex						
Spacegroup $P2_12_12_1$ $a=57.1$ Å, $b=70.3$ Å, $c=184.5$ Å						
Diffraction data						
Data set	$d_{\min}$ (Å)	Independent reflections	Completeness (%)	$I>3\sigma$, (%)	Mean redundancy	R_{merge}^*
Segment 1:actin complex	3.0	15,145	99	86	3.4	0.089
Segment 1(C57N):actin complex	2.5	25,586	96	78	3.3	0.057
Current model						
Number of observations	25,586					
Protein atoms	3,899; actin (2,925); segment 1 (974), 1ATP, 3Ca ²⁺					
Water molecules	52					
$R_{\text{cryst}}^\dagger$ ($d_{\min}$ 2.5 Å)	0.187					
Geometry						
(r.m.s. deviations distances from ideality)(Å)						
Bonds (1–2 neighbours)	0.013 Å					
Angles (1–3 neighbours)	0.051 Å					
Intra planar (1–4 neighbours)	0.060 Å					
Mean B -factor (Actin)	47 Å ²					
Mean B -factor (Segment 1)	37 Å ²					
R.m.s. deviations from mean B -factor						
Bonded main chain atoms	1.6 Å ²					
Bonded side chain atoms	3.0 Å ²					

Diffraction data. Crystals (natSEG1:ACTIN) of segment 1:actin complex (cytoplasmic form of gelsolin residues 26–150), were grown at 20 °C, pH 6.6 in 5–9% polyethylene glycol 6000, 0.15 M NaCl, 0.1 mM ATP, 0.1 mM Ca²⁺, 0.1 mM Mg²⁺, 1 mM NaN₃ (ref. 17). Crystals of cysSEG1:ACTIN, in which segment 1 has an engineered point mutation (residue 57 asparagine replaced by cysteine), were grown under identical conditions except for the addition of 10 mM DTT. The largest natSEG1:ACTIN (0.7 mm × 0.3 mm × 0.1 mm) was used to measure a 3.0 Å data set on station PX 9.6 ($\lambda = 0.88$ Å) at the Daresbury synchrotron using a FAST detector; images were processed to extract reflection intensities using the MADNES³⁹ program. The cysSEG1:ACTIN crystals were obtained from a parallel program to produce point cysteine mutants of segment 1 that might be needed if heavy-atom derivatives were required. Fortunately, the C57N mutant led to larger crystals of the actin complex. One of these (2.5 mm × 1.1 mm × 0.1 mm) was used to measure reflections to 2.5 Å spacings at Daresbury using a MAR Research (J. Hendrix and A. Lentfer) image plate detector (at station PX9.5; $\lambda = 0.90$ Å); intensities were obtained using the MOSFLM package. With the exception of the refinement, crystallographic software used the CCP4 suite (SERC, Daresbury, UK). **Location of actin.** The model derived for the actin part of actin:DNase 1 (ref. 20) was used as a search object in rotation and translation functions operating on the natSEG1:ACTIN diffraction data. These gave significant highest peaks. At this point the 2.5 Å diffraction measurements from the cysSEG1:ACTIN crystals became available and subsequent refinement and model building continued with these alone. **Solvent flattening.** After rigid-body refinement of the molecular replacement solution model against the new data, the map calculated using phases based on this actin model alone was of sufficient quality to show the boundary of the segment 1 molecule but the details of the secondary structural elements were unclear. After 10 rounds of solvent flattening (using a conservative 0.55 solvent fraction, the actual fraction is 0.63) the phases had sufficiently improved to show a helix and a core β -sheet of four strands. **Model building and subsequent refinement.** The segment 1 model was built using a fragment-fitting data base algorithm in the graphics program O⁴⁰. In the first round, 42 residues out of 125 were built. Following a subsequent round of combined energy and X-ray refinement (XPLOR⁴¹) for actin and the partial segment 1 model, another 30 segment 1 residues were built. After six rounds of building and refinement all but the first three amino acids of segment 1 had been located, including two Ca²⁺ ions. After a simulated annealing refinement (XPLOR), assigning individual restrained atomic B-factors, a bulk solvent correction and placing the better order solvent molecules, the agreement factor, R_{cryst} was 0.215. A further two rounds of Hendrickson-Konnert refinement, with a bulk solvent correction, improved R_{cryst} to 0.188 and the atomic model had a good agreement with standard stereochemistry. The model for segment 1 now contains all residues except for the three N-terminal residues. All backbone conformations are in low-energy regions defined by the Ramachandran plot. In the actin model, the first five residues and residues 40–49 (the 'DNase 1 loop') have no convincing electron density; the actin backbone conformations are in allowed regions except for Ser 233 and Ser 235, which lie in a strand of rather broadened density that may have more than one conformation. Atomic coordinates will be deposited in the Brookhaven Protein Data Bank.

* $R_{\text{merge}} = \sum_i |I_i - \langle I \rangle| / \sum_i \langle I \rangle$ where I_i is the intensity measurements for a reflection, and $\langle I \rangle$ is the mean intensity of this reflection.

† $R_{\text{cryst}} = \sum_i |F_o - F_c| / \sum_i F_o$ where F_o and F_c are the observed and calculated structure factors, respectively, and the sum is over all reflections that were measured.

have not changed in the segment 1 complex and they constitute a large fraction of the apolar patch occluded from solvent by segment 1 binding. Bordering the apolar patch, 13 inter-protein hydrogen bonds are made (Fig. 2c). Although four residues have carboxylate side chains, none forms an ion pair. The exclusion of these residues to the rim means that none is buried; this would be energetically unfavourable if the charges were uncompensated. The natural cleft formed between actin subdomains 1 and 3 and the unusual distribution of atom types may be recognized by other actin-binding proteins (indeed, putative myosin head binding residues border the region²⁰).

We have located a Ca^{2+} ion that is bound by ligands from both actin and segment 1 (Fig. 2d). This Ca^{2+} is most probably the one found trapped in gelsolin actin complexes, which greatly enhances the binding affinity^{7,12}. Four ligands, including two

from Asp 109, come from the segment 1 helix and the loop following it. A second carboxylate ligand comes from actin Glu 167'. Few structures have been reported of metal chelation between interacting proteins: preferred binding of Ca^{2+} between subunits occurs in a number of icosahedral viruses, but the nature of these interactions is very different.

A second Ca^{2+} site has been identified in which ligands are from segment 1 alone (Fig. 2e). This calcium is ligated by the two buried carboxylate groups that are conserved in the sequence motifs described earlier. Several lines of evidence suggest that it may be involved in phosphoinositide binding. A segment 1 variant in which 15 C-terminal residues are deleted no longer binds PIP_2 but still binds actin tightly²¹. This region includes residue 145, which ligates the calcium through its carbonyl oxygen. Like annexin V²², a Ca^{2+} -dependent phospho-

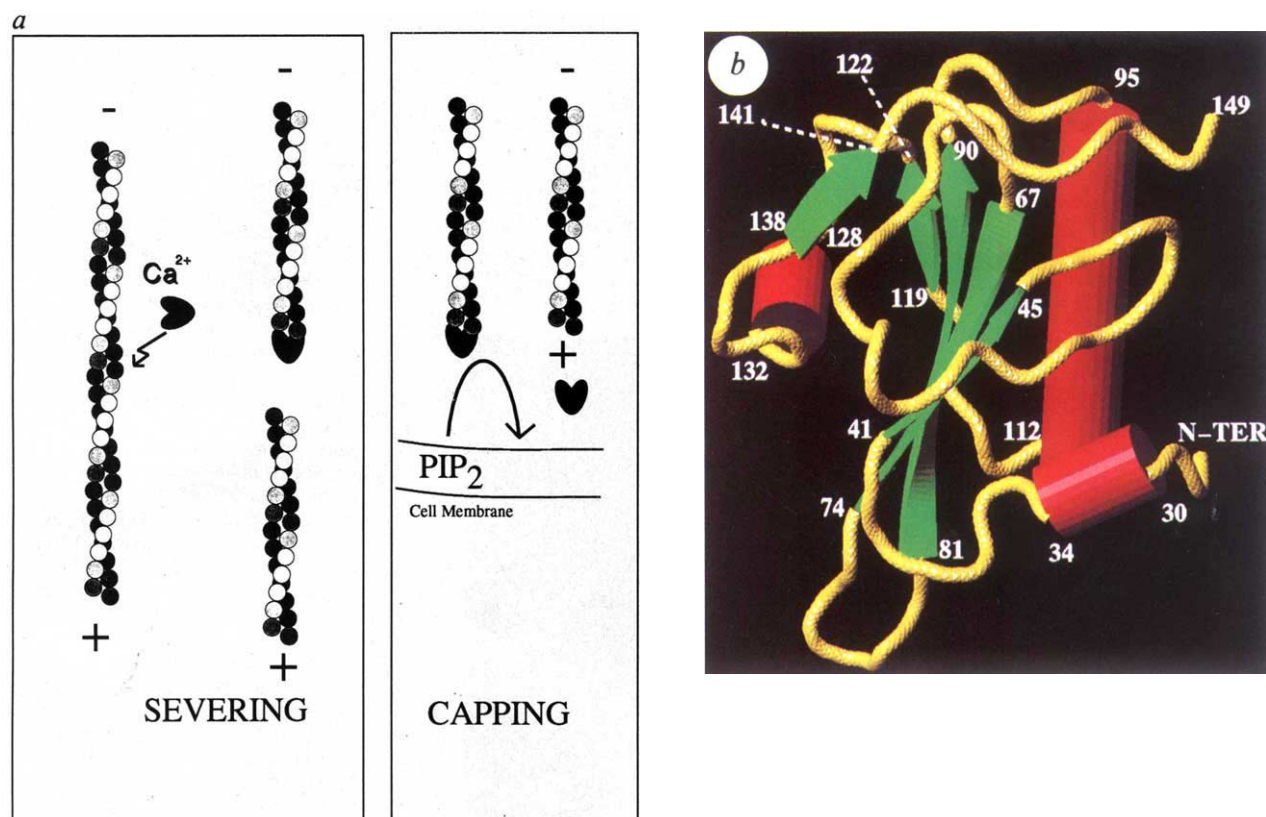
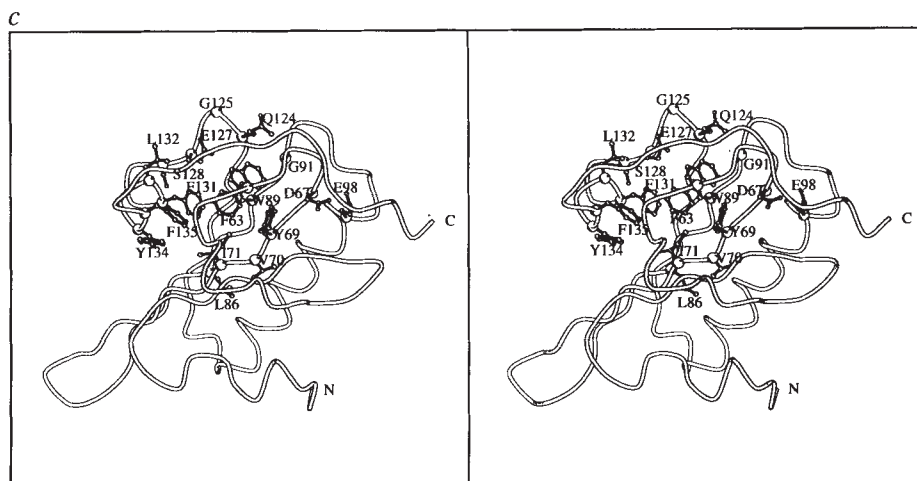


FIG. 1 a, Gelsolin activities. Upon Ca^{2+} -ion activation, at sub-micro-molar concentrations, gelsolin randomly severs actin filaments, tightly binding at the preferred (+) end for monomer addition. PIP_2 micelles can dissociate this tight complex. b, Overview of segment 1 fold. Schematic representation (made by computer program O⁴⁰); α -helices are shown as red cylinders, β -sheets as green arrows and connecting loops as yellow tubes. The packing of helices on either side of a four-stranded sheet is similar to thioredoxin but the order of connection is different. c, Conserved residue positions in segment 1. Stereo view showing segment 1 residues that are conserved in all segments of the gelsolin family. The sequence numbers refer to human plasma gelsolin.



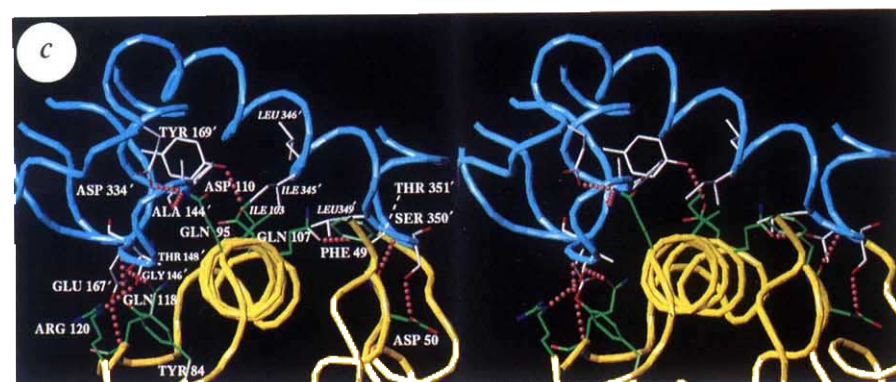
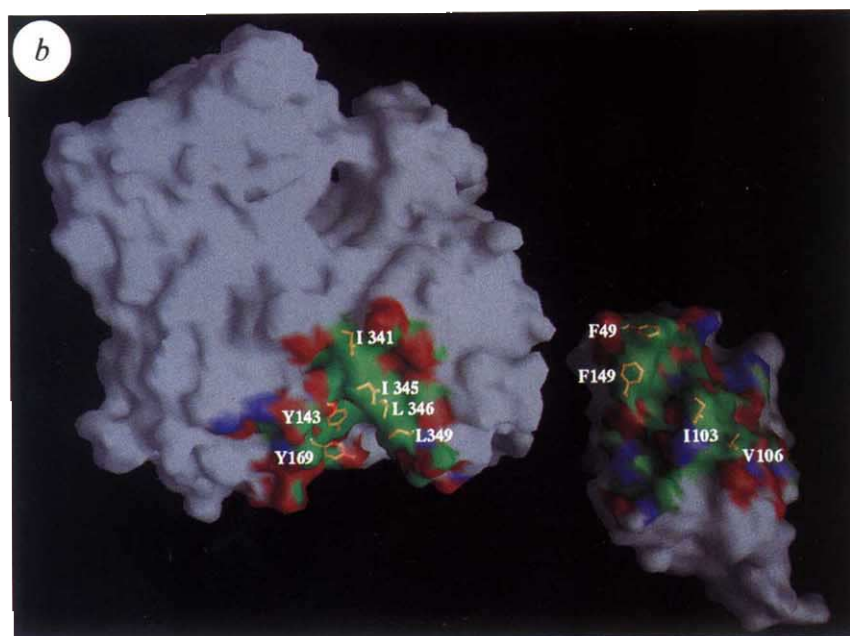
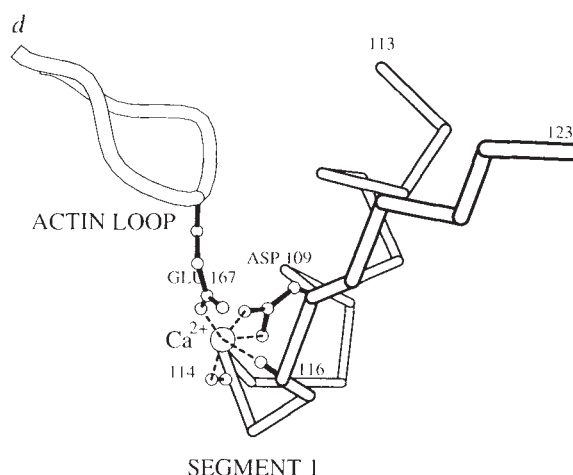
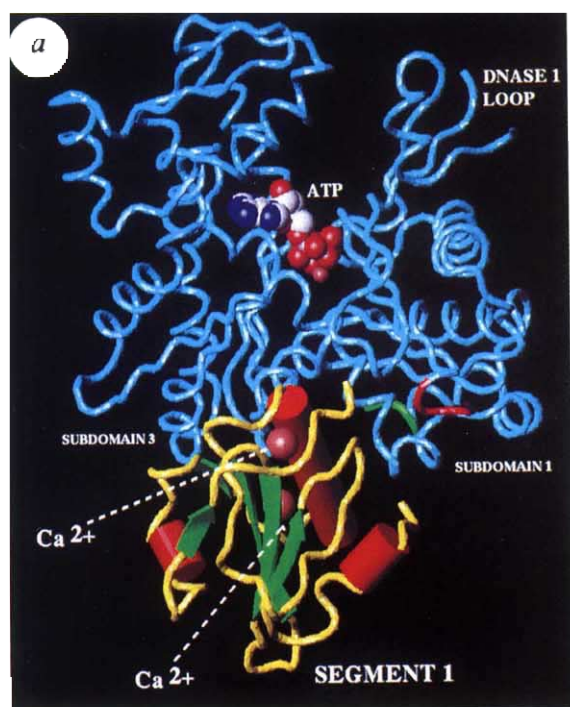
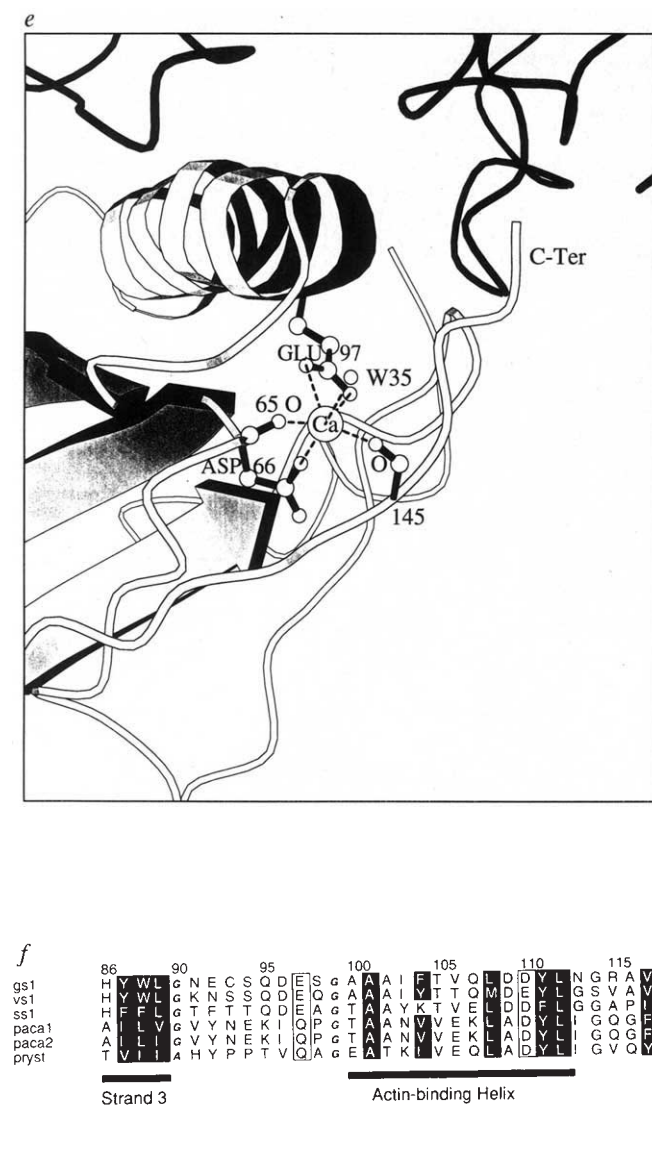


FIG. 2 a, Overview of segment 1:actin complex. Actin (blue tube) in roughly the standard orientation. Two Ca^{2+} ions have been located (pink spheres). The N terminus is in red; the C terminus is green (O^{40}). b, Interface regions in the complex. Molecular surfaces that become buried are coloured: green (carbon), red (oxygen) and blue (nitrogen). Segment 1 has been rotated 180° about the vertical axis and has been translated from the binding site to show the binding face (GRASP⁴²). c, Intermolecular hydrogen bonds. Stereo view of segment 1 (yellow tubes), looking roughly down actin-binding helix, illustrating binding in between actin (blue) subdomains 1 (right) and 3 (left). Possible hydrogen bonds between molecules are shown, as well as the actin residues that pack Ile103 (Tyr169', Ile345', Leu346' and leu349') (O^{40}). d, Intermolecular calcium ion site. Ca^{2+} binds at the C-terminal end of the actin-binding helix in a loop that contributes two main chain carbonyl oxygen ligands (note these are not through alternating residues like E-F hand sites); there is bidentate coordination from aspartate 110 in the helix. A fifth ligand comes from Glu 167' (actin subdomain 3). The Ca^{2+} site is different from any recently surveyed⁴³ (MOLSCRIPT⁴⁴). e, Intramolecular segment 1 calcium ion site. The Ca^{2+} ion bound at a site near the N terminus of the actin-binding helix is shown. Actin subdomains 1 and 3 are top right and left (black). Glu 97 is bidentate; another ligand opposing it in the equatorial plane (defined by both Glu 97 ligands, 65 O and 145 O) would complete the more usual pentagonal-bipyramid co-ordination (MOLSCRIPT⁴⁴). f, Structural context of profilin sequence homology (single-letter amino-acid code). A pattern (probability of random match, $P=0.45 \times 10^{-85}$) was based on the apolar positions (black boxes) with additional requirement for E/Q(97), D/E(110) and G(90/99) (open boxes). A search of the PIR43 sequence data using PIPL program was specific for gelsolin segment 1 sequences and profilins in amoebae and yeast. Some representative matches are shown: profilins paca 1, paca 2 (*Acanthamoeba castellanii*); pryst (*Saccharomyces cerevisiae*) (ALSCRIPT⁴⁵).



lipid-binding protein, the Ca^{2+} site is missing a seventh ligand that would complete the more usual pentagonal-bipyramid coordination. It has been suggested²³ that in annexins the vacant valency directly ligates the phospholipid head group. If PIP_2 binds to segment 1 in a similar way, it would be located at the N terminus of the actin-binding helix in a strategic position to dissociate the tight complex.

Only a few classes of protein complexes have been solved at atomic resolution. Segment 1:actin is similar to other heterologous complexes²⁴ (protease:inhibitor and antibody:antigen). Unlike associations between homologous subunits²⁵, a relatively small surface area ($2,050 \text{ \AA}^2$) is buried and the number of apolar and polar atoms involved is roughly equal. Segment 1 resembles a protease inhibitor, as it binds in a cleft, unlike an antigen:antibody complex in which the contacting faces are flatter; but like the latter, the hydrogen bonds are formed mostly through side-chain and not main-chain atoms.

Does any sequence homology between segment 1 and other actin-binding proteins make sense in the light of the structure? Profilins bind monomeric actin and are similar in sequence length to gelsolin segments²⁶. Profilin sequences from amoebae¹⁵ and yeast share a pattern of apolar residues that underlie the actin-binding site (Fig. 2f). Not only does the shared pattern denote residues forming the buried apolar face of the actin-

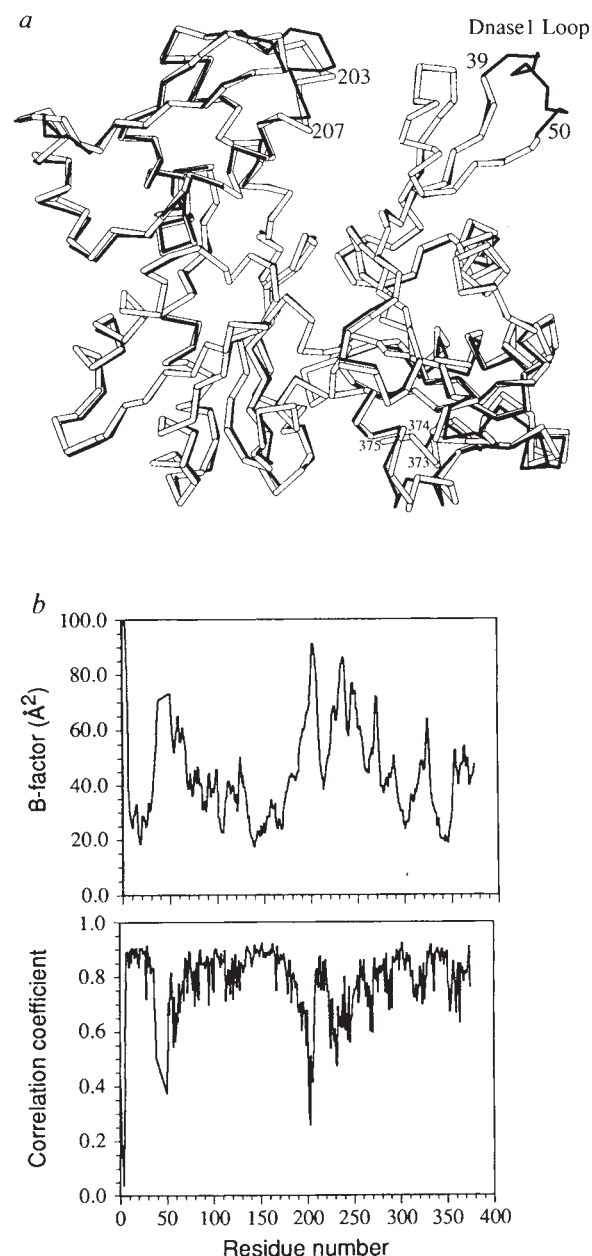


FIG. 3 **a**, Comparison of actin structures. A Ca trace for actin in the complex is shown in open bonds, the superposed actin from DNase 1 structure as closed bonds. Actin main chain atoms in residues 6–39 and 50–372 were superposed using the least-squares algorithm coded in O^{40} . Several features reduce artefacts in the comparison: both types of crystal are grown under similar buffer, precipitant and salt conditions; care had been taken to check for bias during the crystallographic refinement; the crystal contacts are quite different (the segment 1 complex has quite a high solvent content of 63%, 10% higher than actin:DNase 1). Residues 1 to 5 have high crystallographic B -factors in the DNase 1 complex; in segment 1 there is no electron density for them and they are probably highly mobile or disordered. Thus if the N terminus describes a large envelope of conformations, the many cross linking studies mapping to these residues may be less precise than previously thought (MOLSCRIPT⁴⁴). **b**, B -value and real-space correlation coefficient plotted against residue number for actin. The real-space correlation coefficient measures the fit of the observed and calculated electron density, and is calculated for all atoms using O^{40} . There are only two regions where there are dips in the plot. These are the DNase 1 binding loop (39–50) and other former DNase 1 contact in subdomain 4 (201–203). Subdomain 4 has higher B -factors than the average in the actin molecule, similar to the actin:DNase 1 structure, but the real-space correlation coefficient is not markedly worse in this region (181–269).

binding helix, but it also extends on either side to other secondary structural elements in segment 1. Because conserved positions in the segment 1 structure underpin the course of the polypeptide that forms the actin-binding site, the two families may bind actin in similar ways.

Changes in actin structure

It has been suggested that actin is a flexible molecule and that subdomains can change their relative orientations, particularly subdomain 2 (ref. 27). A comparison of the actin structure

reported here with actin:DNase 1 complex should be instructive because segment 1 and DNase 1 bind between different domains at non-overlapping sites. In fact, the two actin structures are remarkably similar with an r.m.s. residual between backbone atoms of only 0.9 Å after superposition (Fig. 3). Subdomains 2 and 4, which both contact DNase 1, maintain their relative orientations in its absence. In addition, out of the four subdomains, subdomain 4 has the highest average crystallographic *B*-factor in both the segment 1 complex and the DNase 1 complex.

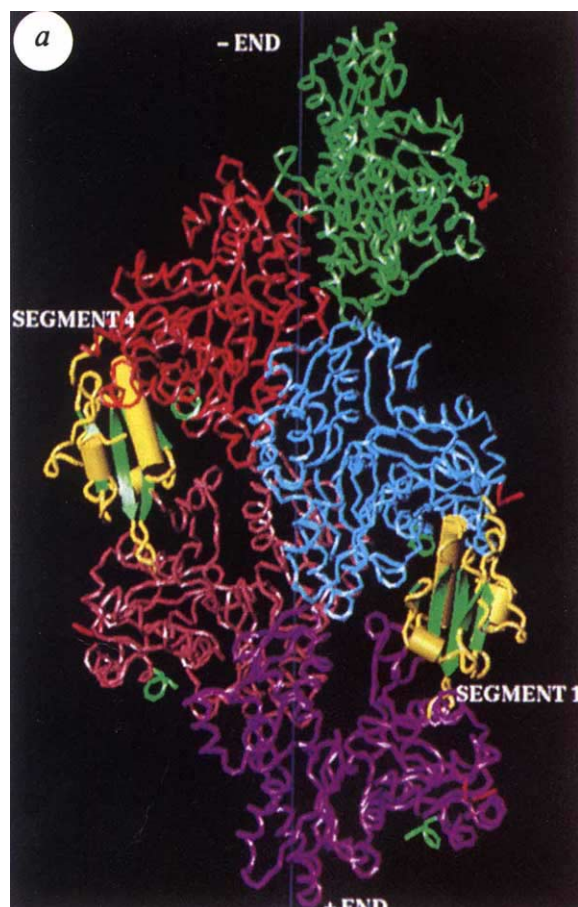
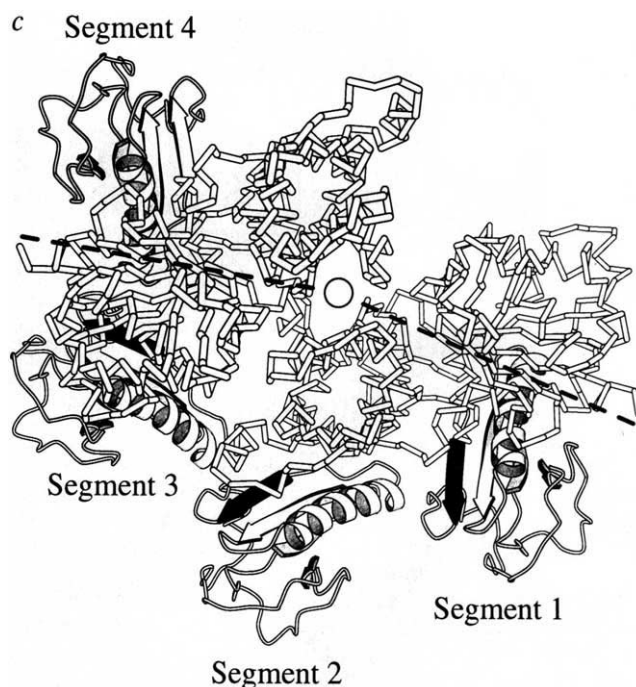
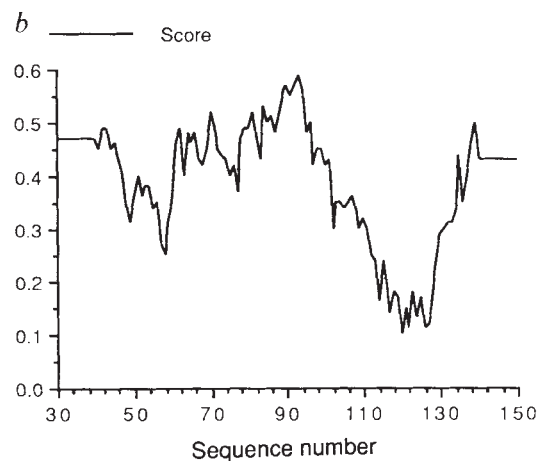


FIG. 4 *a*, Segment 1 bound to fibre-diffraction model for actin-filament. The atomic model of segment 1:actin complex (blue subunit) is shown superposed on a model for the actin filament based on fibre diffraction data. The preferred end for monomer addition is shown at the bottom of the figure (+). The helical symmetry that relates subunits across the filament axis (for example magenta to brown subunits) is a translation by half a subunit height and a rotation of about 166° in an anticlockwise sense about the axis (viewed from the + toward the - end of the axis). The blue subunit is related to the magenta by two such operations and thus lies directly up the axis above the magenta subunit, rotated by 28° in the opposite (clockwise) sense. Thus the filament can be considered as two-start or double-helix in which each strand (magenta:blue:green and brown:red) winds more slowly with an approximate 28° rotation per subunit (O^{40}). *b*, Unusual environment in the segment 1 edge strand. Profile analysis⁴⁶ assigns an environmental class to each position, based on its secondary structure, accessibility to solvent and contact area with polar residues; a score, based on the relative frequency of the amino-acid type in that environment in a database of reliable protein structures, is plotted against segment 1 sequence as a running average (window size 21). The average score for this segment is 42 and is consistent with well refined structures of the same size. But strand 4 (residues 119–122) has a notably lower score (independent of including actin in the calculation). The electron density is clear and crystallographic *B*-factors are low for this region. The environment of strand 4



may change in the context of whole gelsolin and so the strand may be at a segment:segment interface. *c*, Experimental packing for segments around the filament using the smaller angle alternative. Two Ca traces for actin molecules related across the filament axis are shown in open bonds. The view is down the axis from the - end towards the + end. Segment 1 is shown as it is bound in the crystal structure. A putative segment 4 is shown by analogy on the other subunit. Models for segment 2 and 3 are constructed by subsequent rotations by 166°/3 of segment 1 to distribute them evenly around the outside of the filament. This rotation brings segment 2 to a position adjacent to strand 4 (marked in black) of segment 1. Rotation in the opposite sense, about the larger reflex angle, would require long linker loops at once or more segment junctions (MOLSCRIPT⁴⁴).

There are several local differences at the respective binding sites. Two main DNase I contacts are significantly more disordered in the segment 1 complex. Thus in the latter, residues 40–49 have no electron density, whereas in the DNase I complex this loop containing a number of hydrophobic residues forms a parallel β -sheet with DNase I (ref. 20). In the absence of DNase I it is clear that the loop would extend into solvent, so it would not be surprising if the conformation changed; indeed the absence of any electron density for the loop in the segment 1 complex suggests that it is mobile or disordered. Similarly the second DNase I contact region on subdomain 4 (residues 200–207) becomes less well defined in the segment 1 complex, but the main chain density is clear enough to show that the first turn of the α -helix has unwound. These two regions (40–49; 200–207) are the most poorly defined in the actin:segment 1 map, suggesting that they are flexible or disordered. It is perhaps significant that both regions are at a putative subunit contact in the fibre-diffraction model for F-actin¹⁴ (shown in Fig. 4a). There is only one large conformational change at the segment 1 binding site in the complex: Tyr 169' adopts another rotamer conformation that allows it to make a direct hydrogen bond to segment 1. Apart from this, the two actin molecules are remarkably similar at the binding site.

We have located Cys 374 in good electron density in the actin:segment 1 complex. This residue has been a popular target for attaching probes to analyse F-actin structure. It is one of three C-terminal actin residues (373–375) that were proteolytically removed to improve crystal quality of the DNase I complex²⁰. We have located them in a helical conformation, filling a cavity in subdomain 1, which may explain why residues 348–373 are better defined than in actin:DNase I structure. An undecagold label on Cys 374 has been measured at a radius of 27 Å from the filament axis in reconstructed images of F-actin made in vitreous ice²⁸. When coordinates from our model are superimposed on the F-actin structure (Fig. 4a), the sulphur atom is located 28 Å from the filament axis. The agreement is remarkably close, but may be fortuitous because in the segment 1 complex, SγCys 374' is hidden from solvent and must have a different conformation when it is chemically modified.

Actin filament severing

The minimal gelsolin construct that shows full severing activity is segment 1 plus segment 2, the latter containing the F-actin binding site¹³. Segment 1 alone does not sever filaments, but severing activity is obtained when it is linked to an unrelated F-actin binding protein: thus a hybrid between segment 1 and the

F-actin binding domain of α -actinin has severing activity similar to that of gelsolin¹³. We can predict the role played by segment 1 in the severing mechanism based on the F-actin model¹⁴ (Fig. 4a). When the segment 1 structure is bound to an actin subunit in the filament model (Fig. 4a), steric clashes are made only with subdomain 2 of the adjacent (+ended) subunit in the same strand. Thus only one type of actin:actin contact is affected in the cooperatively assembled filament and no contacts are directly broken between strands in the actin double helix. It has been argued that the bonds within each actin strand are stronger than those between strands^{29,30}. It is precisely this stronger bond between subunits that would be disrupted by segment 1 binding. A notable feature is that segment 1 locates tangentially to the side of the filament. It is not inserted to cover the whole interface between the two subunits, which would require a much greater activation energy. Thus a glancing interaction that disrupts the strong critical bonds may be a feature for fragmentation of actin and may be a general feature of proteins that sever other polymers like tubulin³¹.

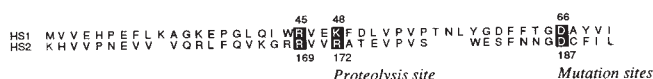
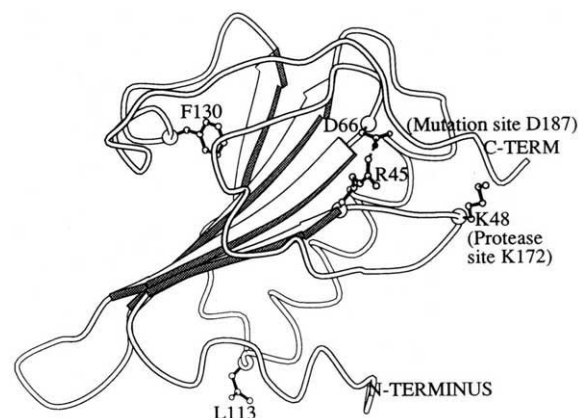
Regions containing weak sequence homology have been reported between gelsolins and actin¹⁶ that might be responsible for severing because they mimicked actin:actin contacts. The structures show that these regions occur in different places in F-actin and segment 1:actin modes. Furthermore, no surface structure in segment 1 seems similar to the actin monomer structure.

Gelsolin binding to F-actin and filament capping

How might all six segments of gelsolin be arranged around an actin filament and cap the +end? Gelsolin binds two actin monomers and acts as a nucleus for polymerization³², which suggests that it binds the two monomers in a filament orientation. Experiments with crosslinked actin dimers have confirmed that gelsolin binds one subunit from each strand, rather than two within a single strand³³. The second actin monomer site in gelsolin is present in segments 4–6: this binds to the same site as segment 1 (ref. 34). Segment 4 and segment 1 are most similar in sequence⁹. They are gene-duplicated equivalents and the actin-binding helix residues are conserved (including the important Ile 103, which is found in no other segment). Furthermore, segment 1 competes with segments 4–6 for the same site on actin³⁴. All this suggests that segment 4 binds an adjacent subunit across the filament axis, in a manner similar to segment 1 (Fig. 4a).

Segments 1 and 4 can be connected around the filament in two non-equivalent ways; either by following the shorter direction around the filament axis that relates subunits in one strand to

FIG. 5 Equivalent residues in segments 1 and 2 involved in gelsolin amyloidosis. The mutation site in segment 2 is Asp 187. The equivalent segment 1 position is Asp 66, which is buried and charge compensated by Arg 45. The equivalent residue in the segment 2 sequence, R 169, is conserved and is close to the newly acquired proteolytic site Arg 172. In the segment 1 structure, the aligned equivalent is Lys 48 which is at the start of a surface loop and the specificity for a trypsin-like protease is conserved. L 113 and F 130 are the segment 1 equivalents of the C termini of the sequenced amyloid peptides (MOLSCRIPT⁴⁴).



the other strand (a 166° rotation component), or by rotation in the opposite sense through the larger reflex angle (194°). Several features inherent in the actin:segment 1 structure favour the smaller angle. First, segment 1 bound to actin is rotated towards subdomain 3 in the same sense as the smaller angle (Figs 2a and 4a). Second, the amino-acid types found in the fourth, edge β -strand of segment 1 are unusual compared to a data base of well refined protein structures (Fig. 4b). This may indicate that the strand is buried in gelsolin; edge β -strand associations are common between protein domains. Packing of a putative segment 2 around the smaller angle would abut this strand. Finally, a simple packing experiment for segments evenly disposed around the shorter angle shows that inter-segmental contacts are possible for domains with the dimensions of segment 1 (Fig. 4c); connection in the opposite sense would leave gaps between two or more segments. All the available evidence suggests that gelsolin segments are likely to be close-packed (and not like 'beads on a string' as in the zinc-finger proteins³⁵).

Amyloidosis Finnish type

Based on the conserved residues in the segment 1 core in the gelsolin family we predict that there will be a common fold in all six segments. If this is true for segment 2, this provides a structural basis to understand the origin of a familial amyloidosis

(Finnish type). In this disease a variant gelsolin, mutated at Asp 187, has acquired a novel site for a trypsin-like protease that leads to deposition of gelsolin fragments as amyloid plaques³⁶. Patients have hyper-elastic skin and suffer from deposits in the cornea and in the cranial nerves (leading to facial muscle weakness). The N-terminal sequence of the isolated amyloid protein is identical to gelsolin, starting at residue 173, 14 residues from the mutation site (Fig. 5). Two variants have been identified in which Asp 187 has changed to either asparagine³⁶ or tyrosine³⁷ in affected families. We infer the three-dimensional relationship of mutation and proteolytic site from the segment 1 structure (Fig. 5). This analysis shows that Asp 187 would form an ion-pair with a buried arginine residue that is close to the cleavage site. Mutation to an uncharged residue would leave the arginine charge uncompensated and consequent local rearrangement of the structure to solvate this charge could easily make the susceptible site more accessible to a protease.

The segment 1 actin model provides a structural basis for understanding the activities of gelsolin which can be tested experimentally. Solving of the crystal structure of whole gelsolin³⁸ is now more tractable with this information. This will reveal the structure of the F-actin binding domain, the spatial relationship between the different actin binding sites and possible synergy between them. □

Received 10 May; accepted 15 July 1993.

- Theriot, J. A. & Mitchison, T. J. *Nature* **352**, 126–131 (1991).
- Stossel, T. P. *J. biol. Chem.* **264**, 18261–18264 (1989).
- Cunningham, C. C., Stossel, T. P. & Kwiatkowski, D. J. *Science* **251**, 1233–1236 (1991).
- DelCastillo, A. R., Vitale, M. L. & Trifaro, J. M. *J. Cell Biol.* **119**, 797–810 (1992).
- Kwiatkowski, D. J. et al. *Nature* **323**, 455–458 (1986).
- Yin, H. L. *Bioessays* **7**, (1988).
- Bryan, J. & Kurth, M. C. *J. biol. Chem.* **259**, 7480–7487 (1984).
- Janmey, P. A. & Stossel, T. P. *J. biol. Chem.* **264**, 4825–4831 (1989).
- Way, M. & Weeds, A. G. *J. molec. Biol.* **203**, 1127–1133 (1988).
- Bryan, J. *J. Cell Biol.* **106**, 1553–1562 (1988).
- Chaponnier, C., Janmey, P. & Yin, H. L. *J. Cell Biol.* **103**, 1473–1481 (1986).
- Way, M., Pope, B., Gooch, J., Hawkins, M. & Weeds, A. G. *EMBO J.* **9**, 4103–4109 (1990).
- Way, M., Pope, B. & Weeds, A. G. *J. Cell Biol.* **119**, 835–842 (1992).
- Holmes, K. C., Popp, D., Gebhard, W. & Kabsch, W. *Nature* **347**, 44–49 (1990).
- Ampe, C. & Vanderkerckhove, J. *EMBO J.* **6**, 4149–4157 (1987).
- Tellam, R. L., Morton, D. J. & Clarke, F. M. *Trends biochem. Sci.* **14**, 130–133 (1989).
- Mannherz, H. G., Gooch, J., Way, M., Weeds, A. G. & McLaughlin, P. J. *J. molec. Biol.* **226**, 899–901 (1992).
- Matsudaira, P. & Janmey, P. A. *Cell* **54**, 139–140 (1988).
- Vanderkerckhove, J. & Weber, K. *J. molec. Biol.* **179**, 391–413 (1984).
- Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. *Nature* **347**, 37–44 (1990).
- Yu, F. X., Sun, H. Q., Janmey, P. A. & Yin, H. L. *J. biol. Chem.* **267**, 14616–14621 (1992).
- Brisson, A., Mosser, G. & Huber, R. *J. molec. Biol.* **220**, 199–203 (1991).
- Huber, R., Schneider, M., Mayr, I., Romisch, J. & Paques, E. P. *FEBS Lett.* **275**, 15–21 (1990).
- Janin, J. & Chothia, C. *J. biol. Chem.* **265**, 16027–16030 (1990).
- Janin, J., Miller, S. & Chothia, C. *J. molec. Biol.* **204**, 155–164 (1988).
- Pollard, T. D. & Cooper, J. A. *Rev. Biochem.* **55**, 987–1035 (1986).
- Orlova, A. & Egelman, E. H. *J. molec. Biol.* **227**, 1043–1053 (1992).
- Milligan, R. A., Whittaker, M. & Safer, D. *Nature* **348**, 217–221 (1990).
- Erickson, H. P. *J. molec. Biol.* **206**, 465–474 (1989).
- Bremer, A. et al. *J. Cell Biol.* **115**, 689–703 (1991).
- Shiina, N., Gotoh, Y. & Nishida, E. *EMBO J.* **11**, 4723–4731 (1992).
- Doi, Y. & Freiden, C. *J. biol. Chem.* **259**, 11868–11875 (1984).
- Doi, Y. *Biochemistry* **31**, 10061–10069 (1992).
- Pope, B., Way, M. & Weeds, A. G. *FEBS Lett.* **280**, 70–74 (1991).
- Nakaseko, Y., Neuhaus, D., Klug, A. & Rhodes, D. *J. molec. Biol.* **228**, 619–636 (1992).
- Maury, C. P. J., Alli, K. & Baumann, M. *FEBS Lett.* **260**, 85–87 (1990).
- DelaChapelle, A. et al. *Nature Genet.* **2**, 157–160 (1992).
- McLaughlin, P. J. & Gooch, J. *FEBS Lett.* **302**, 253–255 (1992).
- Messerschmidt, A. & Pfugrath, J. W. *J. appl. Crystallogr.* **20**, 306–315 (1987).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
- Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Nicholls, A., Bharadwaj, R. & Honig, B. *Biophys. J.* **64**, 166 (1993).
- McPhalen, C. A., Strynadka, N. C. J. & James, M. N. G. *Adv. Protein. Chem.* **42**, 77–144 (1991).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Barton, G. J. *Protein Engng* **6**, 37–40 (1993).
- Luthy, R., Bowie, J. U. & Eisenberg, D. *Nature* **356**, 83–85 (1992).

ACKNOWLEDGEMENTS. We thank W. Kabsch for the coordinates of the actin:DNase 1 complex; the staff at the SERC synchrotron, Daresbury, Warrington, UK; M. Way for engineering the cysteine mutant used in this work; K. Hendrick for his help in porting computer programs; and P. J. Artymiuk, D. Barford, R. A. Cross, R. A. Crowther, P. R. Evans, R. Henderson, A. Leslie, J. Li, J. Kendrick-Jones and P. Stein for discussions. We thank both the Commission of the European Communities and the British-German Academic Research Collaboration Programme for financial support.

A significant contribution to the cosmic X-ray background from sources associated with nearby galaxies

O. Lahav*, A. C. Fabian*, X. Barcons†, E. Boldt‡, J. Butcher§, F. J. Carrera†||, K. Jahoda‡, T. Miyaji‡§, G. C. Stewart§ & R. S. Warwick§

* Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

† Departamento de Física Moderna, Universidad de Cantabria, 39005 Santander, Spain

‡ NASA Goddard Space Flight Center, Code 666, Greenbelt, Maryland 20771, USA

§ X-ray Astronomy Group, Department of Physics, University of Leicester, Leicester LE1 7RH, UK

|| Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

*| Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA

THE origin of the cosmic X-ray background remains a mystery after thirty years of study. The three properties of the background radiation commonly used for tackling this problem—its spectrum, isotropy and resolved component—are well defined by observations, but do not lead to a simple interpretation. A different approach to the problem^{1,2}, in which fluctuations in the unresolved component are cross-correlated with galaxy catalogues, has led to the suggestion² that as much as 60% of the background emission can be explained by a population of X-ray sources similar to present-day optically bright galaxies. Here we point out that such analyses must allow for contributions from X-ray sources which cluster with the galaxies, but do not necessarily have a counterpart in galaxy catalogues. For realistic assumptions about clustering, we obtain a revised limit on the local X-ray emissivity due to sources correlated with nearby galaxies. Extrapolating these results up to a redshift of ~ 5 , we find that a smaller, but still significant, fraction of the X-ray background ($30 \pm 15\%$) can be accounted for by these sources. To explain the residual background emissions, evolution of the source properties and/or a new population of sources at high redshift is required.

The cosmic X-ray background (XRB) is usually discussed in two energy bands. In the soft band below ~ 2 keV an excess is observed in the spectrum³⁻⁵ and much of the emission is resolved into quasars^{4,5}. Above 3 keV, where most of the energy density resides, the spectrum is hard, the sky is isotropic to better than a few per cent, and the total contribution of all known sources appears to account for only 10–20% of the total intensity (see ref. 6 for a review). Source evolution can increase this fraction, but as no sources have the same spectrum as the XRB the interpretation is complicated. Although the origin of the hard XRB remains unresolved, some new insight can be gained from studying statistically the relationship between spatial fluctuations in the XRB intensity and known extragalactic objects. Unlike the other probes of the XRB, which only weakly localize the X-ray sources in space, this cross-correlation probe tells us where the unresolved X-ray sources are, as we know the position and distance to the catalogued galaxies.

We discuss here the cross-correlation function (W_{Xg}) of the unresolved hard XRB with catalogued galaxies in three steps: (1) establishing the detection using two X-ray satellites and several optical and infrared galaxy catalogues; (2) modelling and interpreting the cross-correlation function (W_{Xg}) taking properly into account source clustering—this yields the X-ray volume emissivity in the ‘local universe’ sampled by the galaxy catalogues (at redshift $z \lesssim 0.06$); and (3) extrapolating the local X-ray emissivity to high redshift ($z \sim 5$) to see what fraction of the total XRB can

be explained by such sources.

The results of cross-correlating the positions of galaxies and fluctuations in the hard (2–10 keV) XRB have been published^{1,2}. The first work¹ obtained only a limit on the fraction of the XRB that could be due to sources distributed with normal galaxies. The second work² using the A2 instrument on the High Energy Astronomy Observatory Satellite (HEAO-1), found a positive cross-correlation signal of fluctuations in the XRB intensity $\delta I = (I - \langle I \rangle)$ and in the number of catalogued galaxies $\delta N = (N - \langle N \rangle)$ per beam in the same viewing window (in this case $3^\circ \times 1.5^\circ$): $W_{\text{Xg}} \equiv \langle \delta I \delta N \rangle / \langle I \rangle \langle N \rangle \approx (3 \pm 1) \times 10^{-3}$, where $\langle I \rangle$ and $\langle N \rangle$ are the mean values per beam. This signal can be extrapolated to high redshift (say $z \sim 5$) to give the fraction of the intensity of the hard XRB due to a population of non-evolved sources similar to those correlated with present-day optically bright galaxies. This fraction was originally, but erroneously, given as 13%. It has now been corrected (see erratum in ref. 2) upwards by a factor of 4.5, assuming a Poisson distribution of sources. The interpretation of W_{Xg} strongly depends, however, on whether or not clustering is included. As we show below, the incorporation of clustering effects will revise this extrapolated fraction downwards.

We now generalize the model of the cross-correlation signal for the case in which the sources are clustered. A cross-correlation signal from coinciding (‘zero lag’) X-ray and galaxy viewing windows of finite solid angle may be due to two effects; first, some (or all) the galaxies being the X-ray sources and second, a physical correlation between galaxies and other X-ray sources (which may or may not be the catalogued galaxies). This can be formulated^{7,8} for coinciding cells as a sum of the Poisson term (due to the self-correlation of discrete objects), and a clustering term. In other words, the covariance is increased owing to source clustering. Here we specify the clustering of the X-ray sources with the galaxies in terms of the spatial cross-correlation function of the galaxies with the X-ray emitters as $\xi_{\text{Xg}}(r) = (r/r_0)^{-\gamma}$. The function $\xi_{\text{Xg}}(r)$ is the excess probability (above random) of finding an X-ray source at a distance r from a catalogued galaxy. A power-law correlation function with correlation length r_0 and index γ is a very good empirical fit to the galaxy auto-correlation function⁷. The prediction for the observed W_{Xg} can then be written as

$$W_{\text{Xg}} \langle I \rangle \langle N \rangle \equiv \langle \delta I \delta N \rangle = \eta_{\text{p}} + \eta_{\text{c}} \quad (1)$$

where, assuming for simplicity a square beam of dimension $a \times a$, the purely Poisson term is

$$\eta_{\text{p}} \equiv \rho_{\text{xp}} \frac{1}{4\pi} \int P_{\text{x}}(R) \frac{1}{R^2} d^3 R = \frac{a^2}{4\pi} \rho_{\text{xp}} R_{\text{p}} \quad (2)$$

with $R_{\text{p}} \equiv \int dR P_{\text{x}}(R)$ and $0 \leq P_{\text{x}}(R) \leq 1$ is the estimated probability (that is, selection function) of finding an X-ray emitter in the galaxy catalogue at distance R from us, and the clustering term is

$$\begin{aligned} \eta_{\text{c}} &= \langle n \rangle \rho_{\text{xc}} \int d^3 R_1 \int d^3 R_2 P(R_1) \frac{1}{4\pi R_2^2} \xi_{\text{Xg}}(|\mathbf{R}_2 - \mathbf{R}_1|) \\ &= \frac{a^2}{4\pi} \rho_{\text{xc}} R_{\text{c}} \end{aligned} \quad (3)$$

with $R_{\text{c}} \equiv \langle n \rangle A_{\gamma} r_0^{\gamma} a^{3-\gamma} \int dR R^{3-\gamma} P(R)$, where $\langle n \rangle$ is the mean number density (per unit volume) of the catalogued galaxies (brighter than some L_{min}), $P(R)$ is the estimated selection function of the catalogued galaxies, and A_{γ} is a numerical factor (for $\gamma = 1.8$ and a square beam, $A = 8.3$). The double integral in equation (3) is evaluated assuming the small angle approximation⁷. Note that this clustering term takes into account X-ray sources which are not in the sample of galaxies, but are correlated with them. If none of the catalogued galaxies emits X-rays, then the Poisson term vanishes. The volume emissivity ρ_{xp} is due to the X-ray emitters among the galaxies,

whereas the volume emissivity ρ_{xc} is due to X-ray sources which are correlated with the galaxies. Note that in general both emissivities include contribution from some known discrete X-ray emitters (for example, active galactic nuclei, AGNs), as well as from unresolved sources. We have verified the validity of our formulation, by measuring W_{xg} (assuming a square viewing beam and assigning mock fluxes to particles) in a Cold Dark Matter N -body simulation (S. White, personal communication), for which the spatial correlation function is known, and comparing it with the prediction of equations (1–3).

Clearly, the interpretation of W_{xg} involves many unknowns, and one may consider several extreme cases. For example, the previous analysis² assumed $\xi_{xg}=0$, and derived the volume X-ray emissivity $\rho_{xp} \sim (2 \pm 1) \times 10^{39} h \text{ erg s}^{-1} \text{ Mpc}^{-3}$ (where the Hubble constant $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$). We can see, however, that for fixed W_{xg} and R_p the derived emissivity is smaller if $\xi_{xg} > 0$. Hence source clustering must be taken into account in interpreting the cross-correlation signal.

In what follows we shall assume a model in which the selected X-ray galaxies follow the radial distribution of the optical or infrared (as determined by the Infrared Astronomical Satellite, IRAS) samples, $P_x(R) = P(R)$, and where the X-ray emitters are some (or all) of the galaxies, that is, $\rho_{xp} = \rho_{xc} = \rho_x$ (effectively assuming that all the X-ray sources are galaxies, which are clustered like the catalogued galaxies). In the analyses presented here, known AGNs were not removed before calculating the cross-correlation. Hence the deduced X-ray volume emissivity is due to both identified and unknown X-ray sources. Furthermore, we shall adopt for ξ_{xg} the usual galaxy–galaxy correlation function with $\gamma = 1.8$ and $r_0 = 5h^{-1} \text{ Mpc}$ for optical galaxies⁷ and $\gamma = 1.65$ and $r_0 = 4h^{-1} \text{ Mpc}$ for IRAS galaxies^{9,10}.

To confirm and analyse further the cross-correlation signal we have initiated studies using the UGC¹¹ and IRAS¹² galaxy samples (which overlap in the volume they cover) and X-ray data (in the 4–12 keV band) from another instrument, the Large Area Counter on the Ginga satellite¹³. Taking into account the actual geometry of the Ginga beam (with an effective solid angle of $5.7 \times 10^{-4} \text{ sr}$) and orientation, we found that the cross correlation with the IRAS galaxies brighter than 0.7 Jy is $W_{xg} = 0.015 \pm 0.006$ (the 1σ error bars are extracted by cross-correlating the galaxy sample with random simulations of the XRB. We also find by ‘bootstrap’-resampling² of the flux and number counts that the signal is always positive). Using the IRAS selection function¹⁴ we find $R_p \sim 15h^{-1} \text{ Mpc}$ and $R_c \sim 35h^{-1} \text{ Mpc}$. The cross-correlation with the UGC galaxy samples gives $W_{xg} = 0.012 \pm 0.005$.

Averaging together the optical and IRAS results we find that if the source distribution is Poisson ($R_c=0$) then $\rho_x \sim (6 \pm 3) \times 10^{39} h \text{ erg s}^{-1} \text{ Mpc}^{-3}$. If, however, clustering is assumed (according to the above parameters for ξ_{xg}) then the deduced emissivity is smaller, $\rho_x \sim (1.5 \pm 0.5) \times 10^{39} h \text{ erg s}^{-1} \text{ Mpc}^{-3}$. This value is also confirmed by applying a non-zero lag analysis for the Ginga scans, which allows us to estimate r_0 and ρ_x independently (F.J.C. *et al.*, manuscript in preparation). We have also cross-correlated the HEAO-1 map with galaxies in the IRAS 2 Jy sample¹⁵ at redshift $500 < cz < 8000 \text{ km s}^{-1}$ and 0.7 Jy¹² sample (J.M. *et al.*, manuscript in preparation) and found again a positive cross-correlation signal. Slightly different choices of clustering parameters and selection functions little change the deduced X-ray volume emissivity (F.J.C. *et al.*, T.M. *et al.*, manuscripts in preparation).

The cross-correlation signal hence requires (taking realistic source clustering into account) that the volume X-ray emissivity is $\sim 1 \times 10^{39} h \text{ erg s}^{-1} \text{ Mpc}^{-3}$. If it was due to all bright galaxies, the mean X-ray luminosity per galaxy ($\langle L_x \rangle \sim 10^{40} h^{-2} - 10^{41} h^{-2} \text{ erg s}^{-1}$) would be higher than those determined from studies of individual nearby normal galaxies¹⁶, including the Milky Way. It may be more reasonably explained, however, by assuming that only 10% of the bright galaxies are X-ray sources with $\langle L_x \rangle \sim 10^{41} h^{-2} - 10^{42} h^{-2} \text{ erg s}^{-1}$ in the 2–10 keV band.

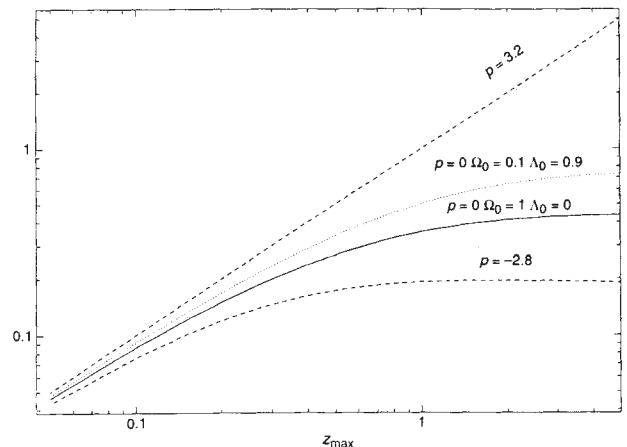


FIG. 1 The effective look-back factor^{8,20} f as a function of maximum redshift $z_{\max}$ for different evolution parameters p . All curves have been computed within a standard cosmological model with $\Omega_0=1$, but for the no-evolution case ($p=0$) the effect of a non-zero cosmological constant Λ_0 in a flat Universe ($\Omega_0 + \Lambda_0 = 1$) is also shown for comparison. An energy spectral index $\alpha = 0.7$ has been assumed. As f is a function of $p - \alpha$, changes in α are equivalent to different values of p . The extrapolated fraction $I(<z_{\max})/\langle I \rangle$ can easily be inferred from this figure by multiplying f by the observed factor $(I_{\text{gal}}/\langle I \rangle)(R_H/R_{\text{eff}})$.

Our estimate for the local emissivity, $\sim 1 \times 10^{39} h \text{ erg s}^{-1} \text{ Mpc}^{-3}$, is comparable to the contribution of bright AGNs, such as Seyfert 1 galaxies, as estimated by Piccinotti *et al.* (ref. 17). This result could be used to constrain the contribution to the hard XRB from low X-ray luminosity AGNs and other classes such as Seyfert 2 galaxies¹⁸ or starburst galaxies¹⁹ in the local universe ($z \lesssim 0.06$).

The significant contribution to the XRB made by sources in the local Universe suggests that the anisotropy of the residual XRB can be used as a probe of structure in the Universe on scales of tens to hundreds of Mpc. If we live in a local region of space which is unusual in overdensity of galaxies and their X-ray emissivity, or if structure within our galaxy samples is strongly correlated with structure at larger distances, then our cross-correlation signal will require further interpretation.

We can now ask what fraction of the XRB can be explained by extrapolating the derived local X-ray emissivity to high redshift $z_{\max}$. The intensity per beam of size a^2 from sources up to $z_{\max}$ is then given by²:

$$I(<z_{\max}) = \frac{a^2}{4\pi} \rho_x \frac{c}{H_0} f(z_{\max}) \quad (4)$$

where $f(z_{\max}) = H_0 \int_0^{z_{\max}} (1+z)^{p-\alpha} (dt/dz) dz$ is the effective look-back factor, which depends upon the spectral index, α (here taken to be independent of z), of the sources, the evolution of their volume emissivity (assumed $\propto (1+z)^p$) and the cosmology^{8,20}. The factor f is plotted against $z_{\max}$ for several representative values of p and cosmological models in Fig. 1. For example, $f(z_{\max}=5, \Omega_0=1, \alpha-p=0.7) = 0.44$ (where Ω is the density parameter). As a fiducial value we shall adopt $f=0.5$. The results given below can easily be scaled for any value of f using Fig. 1.

By combining equations (1) and (4) we can also write the total fraction extrapolated out to $z_{\max}$ as

$$\frac{I(<z_{\max})}{\langle I \rangle} = \left(\frac{I_{\text{gal}}}{\langle I \rangle} \right) \left(\frac{R_H}{R_{\text{eff}}} \right) f \quad (5)$$

where $R_H = c/H_0$ is the Hubble radius, $R_{\text{eff}} \equiv R_p + R_c$, and $I_{\text{gal}} \equiv \langle I \rangle W_{xg} \langle N \rangle$ for a square beam.

Using the cross-correlation of Ginga (taking into account its

true beam shape) with the UGC and IRAS samples we find that $I(< z_{\max})/\langle I \rangle \approx (30 \pm 15)\%$ ($f/0.5$). It encompasses the estimate of $\sim 20\%$ made by Piccinotti *et al.* (ref. 17) on the basis of source counts of bright AGNs. We therefore conclude that the X-ray foreground could be due to AGNs. Modest evolution of AGN can then account for the X-ray background, provided that there is some (not yet fully understood) spectral evolution (for example, ref. 21) because the background has a power-law spectrum with energy index of 0.4, whereas most AGNs at low z have an energy index²² of ~ 0.7 . It is possible that there is another class of source contributing, but it must have a lower local emissivity than AGNs.

Our strong cross-correlation signal is not in conflict with the stringent limits on the auto correlation signals^{13,23,24}. For a euclidean universe, the auto-correlation of the X-ray flux which originates from sources in a shell ($R_{\min}$, $R_{\max}$) in a square beam at 'zero-lag' is

$$W_{xx} \approx (a^2 \langle n \rangle R_{\min}^2 R_{\max}^2)^{-1} + A_7 a^{1-\gamma} / (2-\gamma) (r_0/R_{\max})^\gamma. \quad (6)$$

This illustrates that auto-correlation is a strongly decreasing function of $R_{\max}$ and that it is expected to be much weaker than W_{xx} as it originates from $R_{\max} \gg R_{\text{eff}}$, (that is, the signal is smeared out along the line of sight). Indeed, the analysis of the auto-correlation of the Ginga scans¹³ has shown that an AGN population could produce the entire XRB between redshift 0.05 and 3, if their correlation length is smaller than $9h^{-1}$ Mpc, without exceeding observed limits on the auto-correlation.

In conclusion, the cross-correlation signal indicates that a non-negligible fraction of the XRB background can be produced

by unevolved X-ray sources correlated with optical and IRAS galaxies. However, this is not sufficient to explain the entire XRB and its spectrum, without invoking evolution of the sources, or a new population at high redshift. $\square$

Received 15 February; accepted 29 June 1993.

1. Turner, E. L. & Geller, M. J. *Astrophys. J.* **236**, 1-5 (1980).
2. Jahoda, K., Lahav, O., Mushotzky, R. F. & Boldt, E. A. *Astrophys. J.* **378**, L37-L40 (1991); erratum in *Astrophys. J.* **399**, L107 (1992).
3. Wu, X., Hamilton, T., Helfand, D. J. & Wang, Q. *Astrophys. J.* **379**, 564-575 (1990).
4. Hasinger, G., Schmidt, M. & Trümper, J. *Astr. Astrophys.* **246**, L2-L5 (1991).
5. Shanks, T. *et al. Nature* **353**, 315-320 (1991).
6. Fabian, A. C. & Barcons, X. *A. Rev. Astr. Astrophys.* **30**, 429-456 (1992).
7. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton Univ. Press, 1980).
8. Lahav, O. in *The X-Ray Background* (eds Barcons, X. & Fabian, A. C.) 102-107 (Cambridge Univ. Press, 1992).
9. Lahav, O., Nemiroff, R. J. & Piran, T. *Astrophys. J.* **350**, 119-124 (1990).
10. Saunders, W., Rowan-Robinson, M. & Lawrence, A. *Mon. Not. R. astr. Soc.* **258**, 134-146 (1992).
11. Nilson, P. *Uppsala General Catalogue of Galaxies, Uppsala Astr. Obs. Ann. Vol. 6* (1973).
12. Meurs, E. J. A. & Harmon, R. T. *Astr. Astrophys.* **206**, 53-62 (1989).
13. Carrera, F. J. *et al. Mon. Not. R. astr. Soc.* **260**, 376-384 (1993).
14. Yahil, A., Strauss, M. A., Davis, M. & Huchra, J. P. *Astrophys. J.* **372**, 380-393 (1991).
15. Strauss, M. A. *et al. Astrophys. J. Suppl. Ser.* **83**, 29-63 (1992).
16. Boldt, E. in *The X-Ray Background*, (eds Barcons, X. & Fabian, A. C.) 115-137 (Cambridge Univ. Press, 1992).
17. Piccinotti, G. *et al. Astrophys. J.* **253**, 485-503 (1982).
18. Awaki, H. thesis, Univ. Nagoya (1991).
19. Griffiths, R. E. & Padovani, P. *Astrophys. J.* **360**, 483-489 (1990).
20. Boldt, E. *Phys. Reports* **146**, 215-257 (1987).
21. Schwartz, D. A. & Tucker, W. H. *Astrophys. J.* **332**, 157-162 (1988).
22. Turner, T. J. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **240**, 833-880 (1989).
23. De Zotti, G. *et al. Astrophys. J.* **351**, 22-30 (1990).
24. Jahoda, K. & Mushotzky, R. F. in *The X-Ray Background* (eds Barcons, X. & Fabian, A. C.) 80-96 (Cambridge Univ. Press, 1992).

Ejection of dust from Jupiter's gossamer ring

Douglas P. Hamilton & Joseph A. Burns*

Astronomy Department, and * Theoretical and Applied Mechanics Department, Cornell University, Ithaca, New York 14853, USA

ONE of the most intriguing discoveries of the Ulysses mission so far has been the detection of periodic, collimated streams of high-velocity, submicrometre-sized dust particles emanating from Jupiter^{1,2}. To explain the Ulysses data, Horanyi *et al.* showed³ that electromagnetic forces within Jupiter's magnetosphere can accelerate and eject small dust particles; they proposed a model in which Io is the source of the dust, and the observed periodicity arises from a resonance between the orbital and rotational periods of Io and Jupiter respectively. Here we argue that the masses and velocities of the detected particles are better explained by an origin in Jupiter's gossamer ring. Following their ejection from the magnetosphere, the dust particles are accelerated by the interplanetary magnetic field (IMF). We find that it is the temporal evolution of the IMF which primarily determines the particle trajectories, and hence which particles reach the spacecraft. Our model explains three main features observed in the Ulysses data: fewer streams are detected before closest approach than after, the observed periodicity is closely related to the solar rotation period, and an extremely intense dust stream is detected immediately after closest approach.

We describe first how dust may be ejected from the vicinity of Jupiter (compare ref. 3). When viewed in a non-rotating, jovian-centred reference frame, a charged particle is subject to the so-called co-rotational electric field which arises from Jupiter's spinning magnetic field^{4,5}. The potential associated with this electric field can be readily incorporated into an orbital energy equation if we confine our analysis to nearly equatorial orbits, consider

only the dominant aligned dipolar component of the jovian magnetic field, and assume that the electric charge of a grain remains constant^{6,7}. Because the kinetic energy of escaped particles equals the total energy (kinetic + potential) of grains on initially circular orbits, such particles will be ejected at speed

$$v_x = \left(\frac{2GM_J}{R} \left(L - \frac{1}{2} \right) \right)^{1/2} \quad (1)$$

where G ($=6.67 \times 10^{-8} \text{ cm}^3 \text{ s}^{-2} \text{ g}^{-1}$) is the gravitational constant, M_J ($=1.9 \times 10^{30} \text{ g}$) is the jovian mass, R is the initial distance of the particle from Jupiter, and L is the ratio of the Lorentz force arising from the electric field to Jupiter's gravity⁵. Near Jupiter

$$L \sim 0.0057 \left(\frac{1 \mu\text{m}}{r_g} \right)^2 \left(\frac{\Phi_g}{1 \text{ V}} \right) \left(\frac{1 \text{ g cm}^{-3}}{\rho_g} \right) \quad (2)$$

where Φ_g , r_g , and ρ_g are the surface potential, radius, and mass density of the grain respectively⁵. The charge on circumplanetary grains is determined by a balance between electron, ion, photoelectric, and secondary electron emission currents; in the jovian ring, the latter three currents may be strong enough to induce positive charges. Following refs 3 and 8, we choose $\Phi_g = +3 \text{ V}$ and $\rho = 1 \text{ g cm}^{-3}$ as an illustrative example.

Since v_x in equation (1) must be real, only particles with $L > 1/2$ (which are small and positively-charged) can escape; taking our assumptions for charge and density, $L \geq 1/2$ implies $r_g < r_g^{\text{max}} \sim 0.18 \mu\text{m}$, or $m_g \leq 2.6 \times 10^{-14} \text{ g}$. Particles larger than r_g^{max} are gravitationally dominated and bound to Jupiter; moreover, because L is independent of R , r_g^{max} will not vary with the source location. The tiniest particles have large charge-to-mass ratios and hence spiral (gyrate) tightly around magnetic field lines like electrons and ions do; accordingly these particles are also bound to Jupiter. Particles of intermediate size experience weaker electromagnetic forces and hence have larger gyroradii. Only when gravity is relatively weak and gyroradii are large (for grain radii satisfying $r_g^{\text{min}} < r_g < r_g^{\text{max}}$) do the orbital dynamics

accelerate positively charged particles away from Jupiter (in Table 1 we list $r_g^{\min}$ as determined numerically).

Potential sources of dust in Jupiter's magnetosphere include the main ring⁹ (1.7–1.8 R_J), the gossamer ring¹⁰ (1.8–3.0 R_J), and Io¹¹ (5.9 R_J); dust is thought to be generated in the rings by impacts on parent bodies that slowly abrade away¹², whereas dust that is lofted in Io's volcanic plumes can be swept into the jovian magnetosphere^{3,11}. We now test these sites for their escape characteristics under the assumption that dust begins on initially circular, uninclined keplerian orbits. First, by considering the Jacobi constant, which is conserved in the rotating frame, one can show that particles launched from circular orbits within synchronous orbit ($R_{\text{syn}} = 2.24 R_J$) are unable to escape regardless of their charge histories¹³; thus the main jovian ring can be ruled out as the source of any particles detected by Ulysses.

Table 1 lists the escape characteristics of dust from Io and from two locations in the gossamer ring. Particles from the gossamer ring, especially the denser portion between synchronous orbit and the moon Amalthea (2.5 R_J), are found to escape with maximum speeds similar to those observed by Ulysses² (60 km s⁻¹). Furthermore, the masses of these escaping particles cover the observed two orders of magnitude². By contrast, if volcanos on Io supply the dust streams, then some escaping particles could have speeds of ~300 km s⁻¹, well in excess of those observed by Ulysses, but still within the sensitivity range of its dust detector^{14,15}. In addition, the mass ratio for Io particles is ~8,000, nearly 100 times the measured mass ratio. These results are all robust in that they are independent of the magnitude of the assumed grain charge as long as it remains nearly constant. For Io to be a viable source, an additional mechanism must be invoked to prevent escape of the smallest grains.

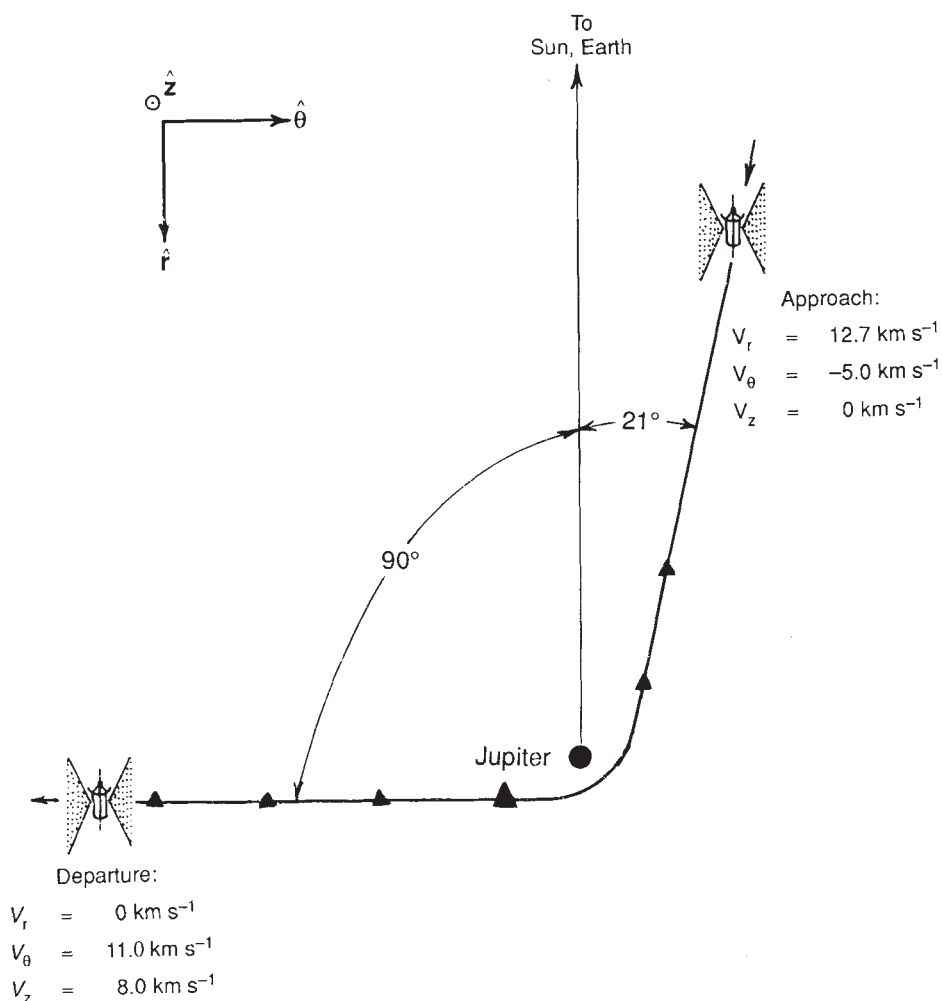
TABLE 1 Properties of particles escaping from Jupiter

Source region	$R(R_J)$	L_{crit}	$r_g^{\min}$ (μm)	$m_g^{\max}/m_g^{\min}$	v_{∞} (km s ⁻¹)
Near synchronous orbit	2.25	4.4	(0.062)	26	<77
Outer gossamer ring	3.0	22	(0.028)	290	<150
Io	5.9	200	(0.009)	8,000	<340

Possible source regions and their distances from Jupiter in jovian radii ($R_J \sim 71,500$ km) are shown, followed by L_{crit} , the numerically-determined force ratio for the smallest escaping grains. The size of these smallest grains ($r_g^{\min}$) is determined from equation (2) using $L = L_{\text{crit}}$, $\rho_g = 1 \text{ g cm}^{-3}$, and $\Phi_g = +3 \text{ V}$. The mass ratio $m_g^{\max}/m_g^{\min}$ is found by solving equation (2) for r_g^3 , using first $L = 1/2$ (which corresponds to the largest escaping grain regardless of source) and then $L = L_{\text{crit}}$. The final column lists numerically determined escape speeds which agree well (to within a few per cent) with those found from equation (1). With the exception of the minimum grain size, all quantities in this table are independent of the assumed grain charge and density.

Particles leave the jovian system along nearly equatorial trajectories because sources are in the equatorial plane and the primary acceleration is radial. Nevertheless, numerical integrations show that the ~10° tilt in the jovian magnetic field causes typical trajectories to bend 10 to 20° out of the plane. These inclinations may allow escaping grains to avoid the densest parts of the Io torus (where grains charge negatively, possibly preventing their escape⁸) by flying above or below it. Later, the distorted field of the middle magnetosphere will augment the bending, but

FIG. 1 The trajectory of the Ulysses spacecraft as seen from Jupiter and projected into the Earth's orbital plane (the ecliptic) which is tilted only a few degrees with respect to Jupiter's equatorial plane. The directions $\hat{r}$, $\hat{\theta}$ and $\hat{z}$ point radially outward, roughly along Jupiter's velocity vector, and toward the ecliptic north respectively. Ulysses, represented by the butterfly-shaped icon, spins around an axis which is directed toward Earth; the shaded triangular wings suggest the spin-averaged directional sensitivity of the dust detector. The components of the spacecraft's velocity relative to Jupiter are listed before and after encounter and, to a good approximation, are valid during all of the stream events. Evenly spaced triangles along the Ulysses trajectory indicate the locations at which dust streams were detected; average stream duration and separation were one day and 28 ± 3 days, respectively². During these events, characteristic particle velocities ranged from 20 to 60 km s⁻¹ while masses varied from 10^{-16} to 10^{-14} g. The stream immediately after closest approach (large triangle) was particularly noteworthy, containing many more particles than all of the other streams combined².



not significantly for swift grains.

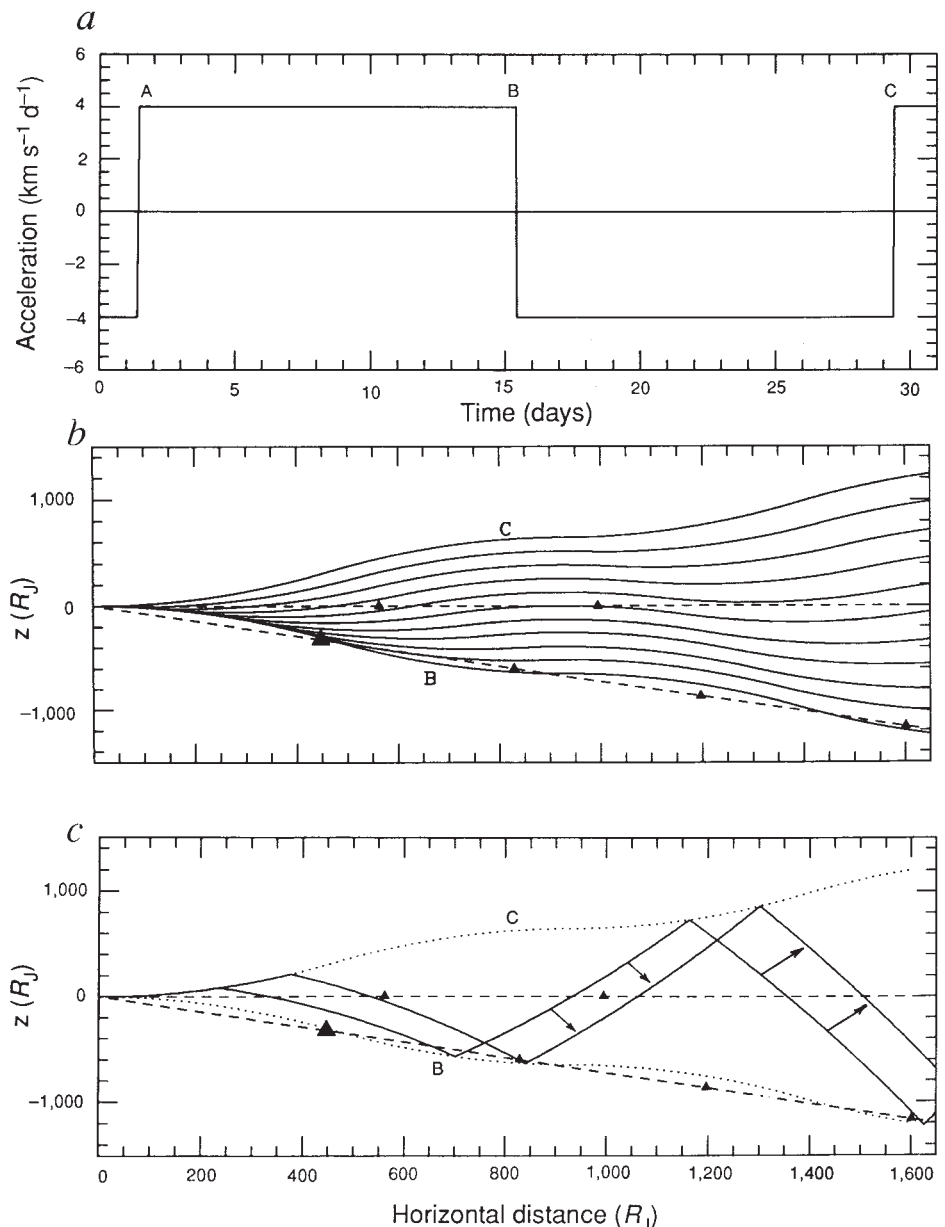
Once in interplanetary space, the motion of a typically-charged grain is dominated by Lorentz forces which greatly exceed solar gravity². On average, the interplanetary magnetic field at Jupiter lies in the ecliptic, counterclockwise from (and within 10° of) the $\hat{\theta}$ direction (Fig. 1)¹⁶. Thus, because the solar wind convects the magnetic field with velocity $v_{sw}\hat{r}$, the induced electric field $\mathbf{E} = -v_{sw}\hat{r} \times \mathbf{B}$, and hence the vertical acceleration, points primarily in the $\pm\hat{z}$ direction (toward or away from the ecliptic plane). The magnitude of this acceleration varies because both the solar wind's velocity and especially its embedded $\mathbf{B}$ field change with a roughly 26–29-day periodicity imposed by solar rotation; stochastic variations further complicate this simple picture¹⁷. Nevertheless, throughout the interplanetary portion of the trajectory of a dust grain, vertical ($\hat{z}$) accelerations overwhelm horizontal ones indicating that the ecliptic component of the grain's velocity remains essentially constant.

We now demonstrate how a periodic vertical acceleration may lead to episodic detections of dust. We represent the acceleration by successive intervals of constant strengths and consider both a symmetric profile (Fig. 2a) and an asymmetric one (Fig. 3a). Vertical displacements are found by integrating the acceleration

twice; these solutions consist of linked parabolic segments (Figs 2b and 3b). In both cases, the exact three-dimensional distribution of all ejected dust particles is complicated, depending on the range of dust particle positions and velocities upon escape from the jovian magnetosphere, as well as the detailed history of the solar wind speed and magnetic field strength. Nevertheless, we can still draw some qualitative conclusions.

According to our interpretation, the periodicity of the Ulysses detections stems from the periodicity of the vertical acceleration which causes the paths of particles to intersect the spacecraft trajectory only at certain times. As time progresses in Figs 2c and 3c, each element of the dust distribution moves to the right along its own streamline at the ejection speed; thus a stationary point near streamline B will receive dust only once every solar cycle (~28 days). Stationary points near the ecliptic plane should experience subharmonic streams at roughly twice this rate (every 14 days). These periods will differ somewhat from those measured by Ulysses because Ulysses moves to a new position and a new streamline between detections. Owing to the directional sensitivity of the dust detector, the subharmonic streams predicted in the ecliptic could be weak; nevertheless, they should be sought when the Ulysses data is further processed.

FIG. 2 Motion of charged dust grains relative to Jupiter. Grains are accelerated by a time-varying symmetric interplanetary magnetic field after being launched from Jupiter in the ecliptic plane with velocities of 40 km s^{-1} . a, The acceleration experienced by a charged grain is constant ($4 \text{ km s}^{-1} \text{ d}^{-1}$), but changes sign every 14 days (roughly half of a solar rotation period); its magnitude corresponds to a $0.1 \mu\text{m}$ grain and solar wind parameters of $B \sim 5 \text{ nT}$ and $v_{sw} \sim 600 \text{ km s}^{-1}$ (ref. 2). b, Streamlines for particles launched between times B and C in a (that is at various times when the initial acceleration is downward). The vertical ($\hat{z}$) direction is defined in Fig. 1. Jupiter is at (0, 0), the horizontal dashed line is Ulysses' incoming trajectory and the diagonally descending dashed line represents its outgoing path. The solid triangles show the locations of the detected dust streams with the largest signifying the single intense burst. Note that the bottommost streamline corresponds to particles launched at time B in a, whereas the topmost is for those starting at time C; the undulations visible in each streamline are caused by changes in the sign of the acceleration. Particles launched between times A and B follow streamlines that are the mirror image of those shown here; in particular the dense concentration of streamlines is found above the ecliptic plane. c, Streamlines B and C (dotted) and the Ulysses path (dashed) are as in b. The solid lines oscillating back and forth through $z=0$ are snapshots of a pair of dust distributions for grains continuously injected over 50 days. The two distributions are separated by about 4 days with the arrows indicating the motion in the intervening time.



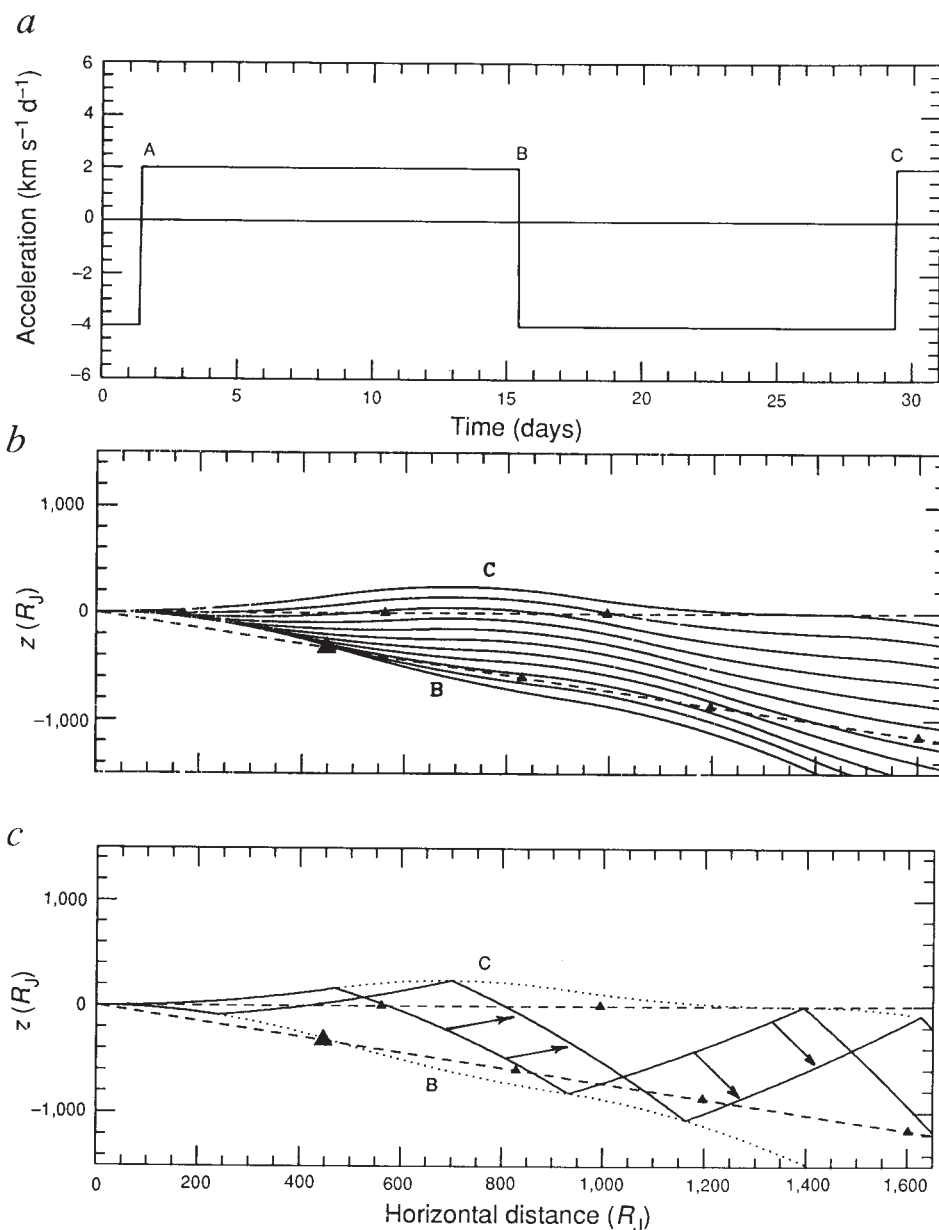


FIG. 3 Motion of a dust grain relative to Jupiter in an asymmetric model of the interplanetary magnetic field; symbols and axes are the same as the corresponding panels (a, b, c) of Fig. 2. *a*, The acceleration experienced by a charged grain in the time-varying asymmetric interplanetary magnetic field. *b*, Because of the asymmetry, the average acceleration felt by a grain over a long period is negative; hence streamlines are swept downward eventually leaving the ecliptic plane free of dust. *c*, The two dust distributions shown here are separated by 7 days; both occur at slightly later times than the distributions in Fig. 2c.

Differences in the number of streams detected in and out of the ecliptic plane can arise from asymmetry in the vertical acceleration (Fig. 3), which itself has many causes: corotating interaction regions, configuration of the solar current sheet, distribution of sunspots and coronal holes and so on^{17,18}. Asymmetry imparts a non-zero vertical acceleration which sweeps particles away from the ecliptic plane. This may explain why only two dust streams were detected before closest approach (when the spacecraft was in the ecliptic plane), whereas four were sensed afterwards (with the spacecraft descending beneath the plane). The detection geometry, which was more favourable after closest approach than before, may also be important (Fig. 1).

Streamlines are clearly most concentrated at the lower left in Figs 2b and 3b; half a cycle later these enhancements are found at the upper left. Streamline clustering occurs because particles launched at different times will follow the same parabolic trajectory through space as long as the acceleration remains pointed in one direction. Furthermore, when $L \gg 1/2$ the ejection mechanism and the subsequent vertical acceleration are both wholly electromagnetic and even particles of different sizes will initially follow the identical path albeit with different speeds. These dense

concentrations of streamlines never cross the ecliptic and disperse once the acceleration changes sign. Our model therefore predicts that intense streams should be detected only near Jupiter and only out of the ecliptic plane, in agreement with the timing and position of the single intense stream detected ~30 days after closest approach (Figs 2 and 3).

We conclude that solar modulations of the interplanetary magnetic field are primarily responsible for the periodicity, distribution, and relative intensities of stream events detected by the Ulysses spacecraft and that the gossamer ring is their most likely source. These results are not critically dependent on the magnitude of the assumed grain charge. $\square$

Received 25 May; accepted 14 June 1993.

1. Grün, E. *et al.* *Science* **257**, 1550–1552 (1992).
2. Grün, E. *et al.* *Nature* **362**, 428–430 (1993).
3. Horanyi, M., Morfill, G. & Grün, E. *Nature* **363**, 144–146 (1993).
4. Burns, J. A. & Schaffer, L. E. *Nature* **337**, 340–343 (1989).
5. Hamilton, D. P. *Icarus* **101**, 244–264 (1993).
6. Schaffer, L. E. *thesis* Cornell Univ. (1989).
7. Horanyi, M. & Burns, J. A. *J. geophys. Res.* **96**, 19283–19289 (1991).
8. Horanyi, M., Morfill, G. & Grün, E. *J. geophys. Res.* (submitted).
9. Showalter, M. R., Burns, J. A., Cuzzi, J. N. & Pollack J. B. *Icarus* **69**, 458–498 (1987).
10. Showalter, M. R., Burns, J. A., Cuzzi, J. N. & Pollack J. B. *Nature* **316**, 526–528 (1985).

11. Johnson, T. V., Morfill, G. & Grün, E. *Geophys. Res. Lett.* **7**, 305–308 (1980).
12. Burns, J. A., Showalter, M. R. & Morfill, G. E. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 200–272 (Univ. of Arizona, Tucson, 1984).
13. Schaffer, L. E. & Burns, J. A. *J. geophys. Res.* (submitted).
14. Grün, E. *et al. Space Sci. Rev.* **60**, 317–340 (1992).
15. Grün, E. *et al. Astr. Astrophys. Suppl. Ser.* **92**, 411–423 (1992).
16. Fisk, L. A. in *The Ancient Sun* (eds Pepin, R. O., Eddy, J. A. & Merrill, R. B.) 103–118 (Pergamon, New York, 1980).

17. Smith, E. J. in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) 175–201 (Univ. of Arizona, Tucson, 1991).
18. Smith, E. J., Slavin, J. A. & Thomas, B. T. in *The Sun and the Heliosphere in Three Dimensions* (ed. Marsden, R. G.) 267–274 (Reidel, Dordrecht, 1986).

ACKNOWLEDGEMENTS. We thank M. Horanyi for discussion, E. Grün for sharing data, and V. Chang, I. Mosqueira and P. Nicholson for critically reading the manuscript, and C. Murray for comments.

Activation of a carbon–carbon bond in solution by transition-metal insertion

Michael Gozin, Alexander Welsman,
Yehoshua Ben-David & David Milstein*

Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

* To whom correspondence should be addressed

CLEAVAGE of carbon–carbon bonds by transition-metal-containing heterogeneous catalysts forms the basis of one of the most important industrial processes, the refining of petroleum to chemicals and fuels¹. But the generally low product selectivity observed with heterogeneous systems is a significant drawback. For this reason, much effort has been devoted to developing transition-metal complexes that can be inserted into C–C bonds in homogeneous media, as such catalysts can operate under mild, easily controlled conditions and might offer high selectivity and reactivity. Metal insertion into C–H bonds is well known^{2,3–5}, but, except in a few special cases^{2,6–12}, C–C bonds are generally unreactive towards insertion of transition metals in solution. Here we report the selective activation of a simple C–C bond by a mononuclear rhodium complex in a neutral homogeneous medium (tetrahydrofuran solution). We are able to effect Rh insertion into a C–C bond in a diphosphinoxyethylene, in which this bond is favourably oriented towards the transition-metal centre. The competing reaction of insertion into C–H is suppressed by the use of an overpressure of H₂. We suggest that this approach might lead to a general strategy for C–C activation in homogeneous systems.

Carbon–carbon bonds are generally unreactive towards the insertion of soluble metal complexes, except when activated by strain^{2,6}, when there is a drive to aromatic character in pre-aromatic systems^{7–10}, or in the presence of a carbonyl group^{11,12}. β -Alkyl transfer in highly Lewis-acidic complexes has been

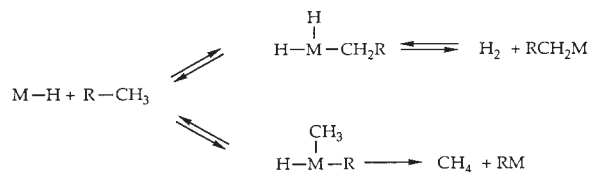


FIG. 1 Hypothesis for C–C activation under hydrogen. (R is an alkyl or aryl group.)

reported^{13,14}. To investigate the possibility of transition-metal insertion into a C–C bond in solution, we planned to use a metal hydride complex, which (if capable of insertion into C–C) would eliminate methane and thereby render the process thermodynamically feasible. The expected competing reaction, the kinetically favourable insertion into C–H, would probably be reversible even if hydrogen is eliminated and might be pushed backwards under hydrogen (Fig. 1).

α,α' -Diphosphino-*m*-xylenes undergo C–H activation forming bis-chelated complexes^{15–17}. Probing for the possibility of directed C–C cleavage, we prepared the new phosphine, shown as **1** in Fig. 2. Reaction of **1** with $\text{HRh}(\text{P}(\text{C}_6\text{H}_5)_3)_4$ in a tetrahydrofuran (THF) solution at 25 °C results in exclusive C–H activation to yield quantitatively the thermally stable Rh(I) complex **2**, with liberation of hydrogen. The pure complex was fully characterized by ¹H, ¹³C and ³¹P NMR and its structure was confirmed by a single-crystal X-ray structural study (Fig. 3), which shows that (1) the metal-bound Ar–C(7) bond (where Ar = aryl) is not weakened relative to the other Ar–C bonds in this molecule, (2) there is no distortion from aromaticity, the aromatic ring and the carbon atoms bound to it retaining planarity, (3) to maintain a favourable square-planar arrangement around the rhodium atom, a relatively weak Rh–CH₂Ar bond is formed as reflected by the distorted C(1)–C(7)–Rh(I) angle of 92.8 (4)°.

Remarkably, heating of **2** at 90 °C under hydrogen at a pressure of 80 p.s.i. results in quantitative C–C cleavage, yielding

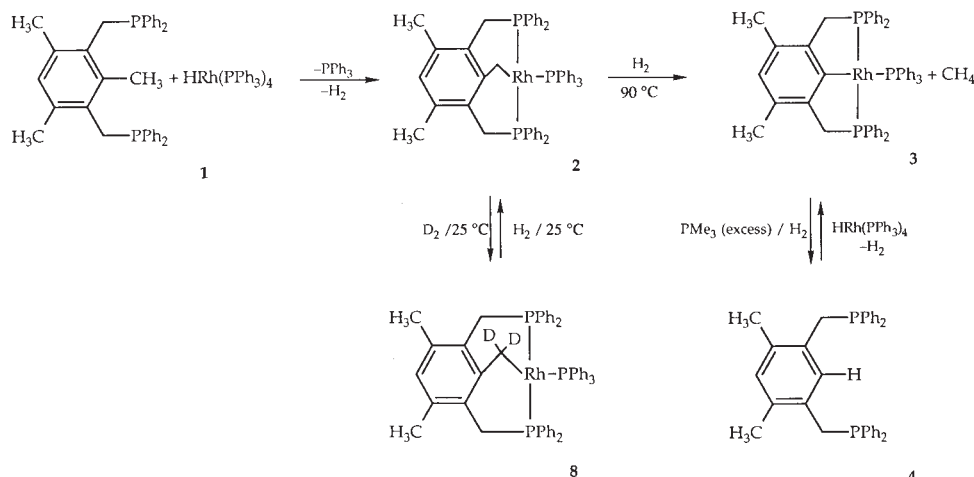


FIG. 2 Experiments demonstrating competitive C–H and C–C activation, resulting in selective C–C hydrogenolysis. (Ph is a phenyl (C₆H₅) group, Me is a methyl (CH₃) group.)

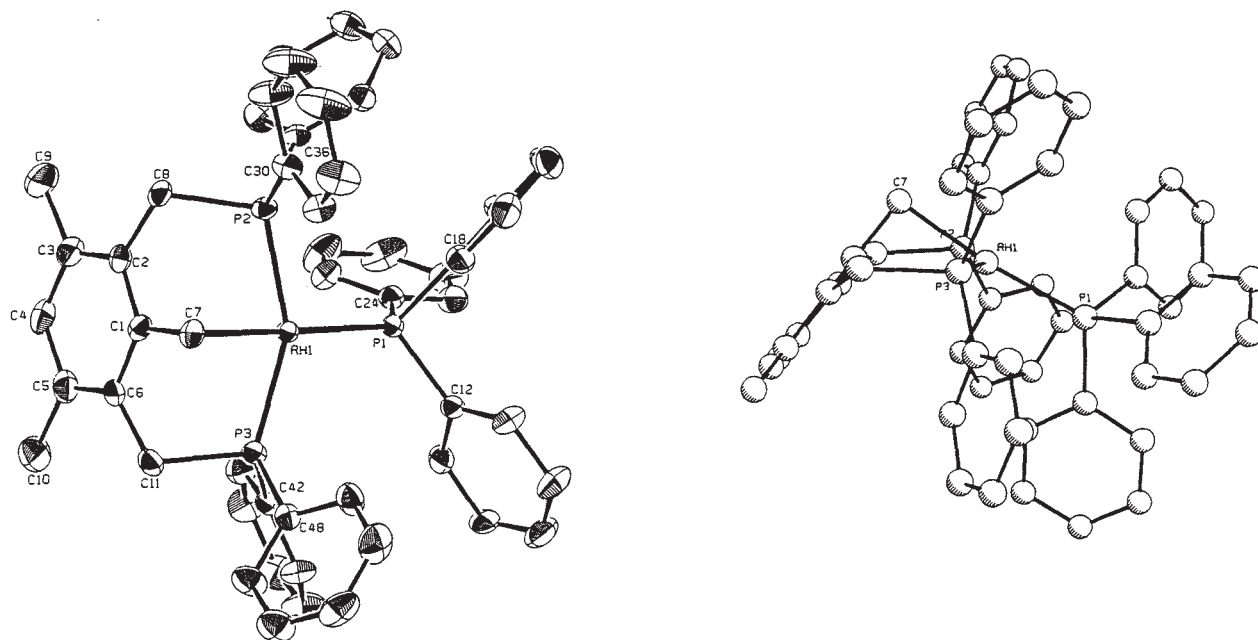


FIG. 3 Structure of a molecule of **2** (from Fig. 2), viewed from two different angles.

complex **3**. Analysis of the gas phase reveals quantitative formation of methane. The structure of **3** is unambiguous, based on ^1H , ^{31}P and ^{13}C NMR measurements, and it is further confirmed by its independent synthesis from reaction of de-methylated ligand **4** with $\text{HRh}(\text{P}(\text{C}_6\text{H}_5)_3)_4$. Treatment of a solution of complex **3** with excess trimethylphosphine under hydrogen (20 p.s.i.) at 70°C yields the phosphine **4**, resulting in overall specific hydrogenolysis of one of the methyl groups (Fig. 2).

A plausible mechanism for this unprecedented transformation is presented in Fig. 4. Although H_2 addition to **2** is evidently thermodynamically unfavourable, the set of equilibria between **2** and intermediates **5** and **6** is well supported. Thus, treatment of **2** with deuterium (20 p.s.i.) at 25°C resulted in complete H/D exchange of the hydrogens on the Rh-bound benzyl to yield **8**, whereas the other benzylic hydrogens remain unaffected.

Insertion of the $\text{HRh}(\text{I})$ centre into the C–C bond in **6** yields the postulated methyl hydride intermediate **7**, which liberates methane irreversibly to give the demethylated **3**. C–H elimination from $\text{Rh}(\text{III})$ phosphine complexes, involving a pentacoord-

inate intermediate has been demonstrated¹⁸.

Regarding the mechanism of the insertion into C–C (step **6** $\rightarrow$ **7**) a simple one-step process proceeding through a 3-centred transition state, **9**, is not the only possibility. Complex **6** may be in equilibrium with an η^2 -arene intermediate **10**, which undergoes a 1,2-methyl migration from the aryl carbon to the metal centre. The relatively short Rh–C(1) distance of 2.66 Å in **2** may suggest some interaction between the rhodium centre and the aromatic ring, although we do not have ^{13}C NMR evidence to support such an interaction.

It is relevant that van Koten and co-workers have isolated a σ -metal-substituted arenonium cation from the reaction of a σ -aryldiamino platinum(II) complex with methyl iodide¹⁹.

There are numerous examples of homogeneous metal insertions into C–H bonds^{2,3,5}, which may be attributed to a higher kinetic barrier for C–C activation due to lesser accessibility of these bonds and the higher directionality of M–C as compared with M–H bonds^{20,21}. Thermodynamic constraints, which need compensation from strain relief or aromatization, are probably

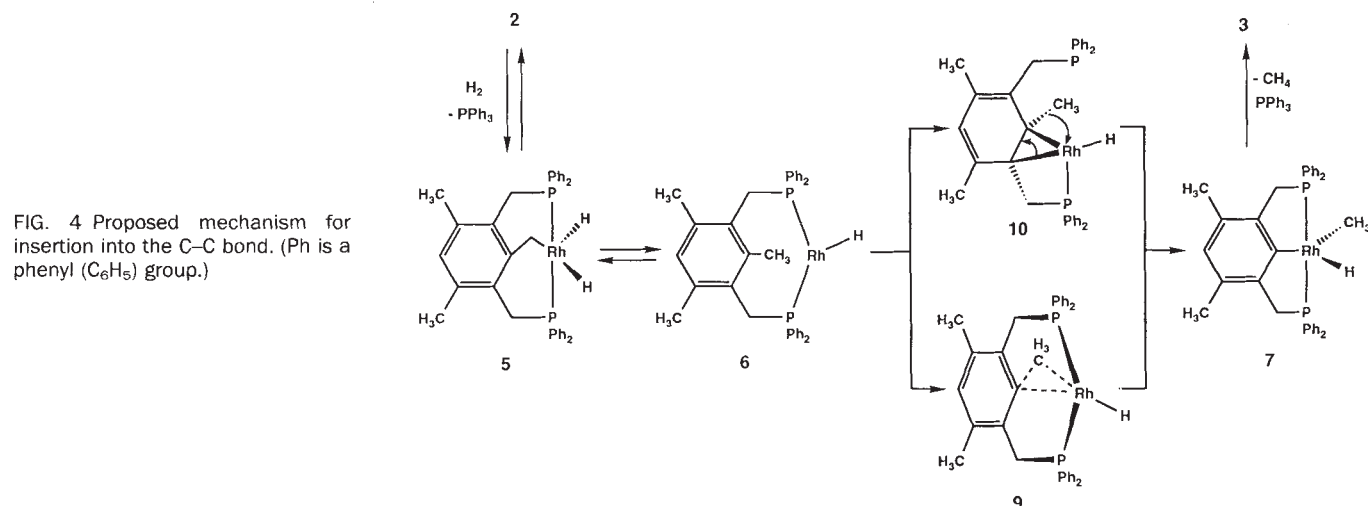


FIG. 4 Proposed mechanism for insertion into the C–C bond. (Ph is a phenyl (C_6H_5) group.)

less important. Indeed, several stable dialkyl complexes are known. Our results clearly show that C–H activation is kinetically favoured over C–C activation, in agreement with theoretical calculations^{20,21}. This kinetic barrier for C–C cleavage is overcome here by relatively mild heating. As **3** does not react with methane, the overall C–C activation process is irreversible and is thermodynamically more favourable than the overall C–H activation sequence, although both processes are clearly possible. We cannot rigorously compare the single insertion steps at this stage. It may seem that insertion into ArCH₂–H and generation of Rh–H would be thermodynamically more favourable than insertion into Ar–CH₃ (bond dissociation energies (BDE) of C₆H₅CH₂–H and C₆H₅CH₃ are 88.0±1 and 101.8±2 kcal mol⁻¹ respectively²²). We have to remember, however, that Ar–Rh is considerably stronger than alkyl–Rh^{23,24}, a relationship that is particularly applicable for **2** (see X-ray structure), and that the Rh–H bond in **5** is relatively weak (spontaneous H₂ elimination from **5** indicates a BDE of Rh–H < 52 kcal mol⁻¹). Thus, metal insertion into C–C may actually prevail even in this case, even though the C–C bond is probably stronger than the C–H bond. Note that bond dissociation energies of C–H bonds do not correlate well with their relative reactivities in C–H oxidative addition². Also, an *sp*²–*sp*² C–C bond may be more reactive than an *sp*³–*sp*³ C–C bond, in line with the typically higher oxidative addition reactivity of aryl–H relative to alkyl–H.

It should be possible to find other systems in which 'precoordination' in the substrate enhances metal insertion into a C–C bond. Moreover, although C–C activation is helped here by metal precoordination, we believe that it need not be a prerequisite. This work provides hope that homogeneous activation and

thus selective transformation of C–C bonds by metal insertion may be achieved quite generally. □

Received 12 May; accepted 7 July 1993.

1. Sinnett, J. H. in *Catalysis: Science and Technology* Vol. 1 (eds Anderson, J. R. & Boudart, M.) 257–300 (Springer, New York, 1981).
2. Crabtree, R. H. *Chem. Rev.* **85**, 245–269 (1985).
3. Bergman, R. G. *Science* **223**, 902–908 (1984).
4. Shilov, A. E. *Activation of Saturated Hydrocarbons by Transition Metal Complexes* (Reidal, Boston, 1984).
5. Maugh, T. H. *Science* **220**, 1262–1263 (1983).
6. Periana, R. A. & Bergman, R. G. *J. Am. chem. Soc.* **108**, 7346–7355 (1986).
7. Crabtree, R. H., Dion, R. P., Gibboni, D. J., McGrath, D. V. & Holt, E. M. *J. Am. chem. Soc.* **108**, 7222–7227 (1986).
8. Kang, J. W., Moseley, R. & Maitlis, D. M. *J. Am. chem. Soc.* **91**, 5970–5977 (1969).
9. Benfield, F. W. C. & Green, M. L. H. *J. chem. Soc. Dalton Trans.* 1324–1331 (1974).
10. Eilbracht, P. *Chem. Ber.* **113**, 542–554 (1980).
11. Suggs, J. W. & Jun, C.-H. *J. Am. chem. Soc.* **106**, 3054–3056 (1984).
12. Hartwig, J. F., Anderson, R. A. & Bergman, R. G. *J. Am. chem. Soc.* **111**, 2717–2719 (1989).
13. Watson, P. L. & Roe, D. C. *J. Am. chem. Soc.* **104**, 6471–6473 (1982).
14. Bunel, E., Burger, B. J. & Bercaw, J. E. *J. Am. chem. Soc.* **110**, 976–978 (1988).
15. Moulton, C. J. & Shaw, B. L. *J. chem. Soc. Dalton Trans.* 1020–1024 (1976).
16. Rimm, H. & Venanzi, L. M. *J. Organomet. Chem.* **259**, C6–C7 (1983).
17. Nemeš, S., Jensen, C., Binamira-Soriaga, E. & Kaska, W. C. *Organometallics* **2**, 1442–1447 (1983).
18. Milstein, D. *Accs. chem. Res.* **17**, 221–226 (1984).
19. Grove, D. M. et al. *J. Am. chem. Soc.* **104**, 6609–6616 (1982).
20. Blomberg, M. R. A., Siegbahn, P. E. M., Nagashima, U. & Wennerberg, J. J. *J. Am. chem. Soc.* **113**, 424–433 (1991).
21. Low, J. J. & Goddard, W. A. *Organometallics* **5**, 609–622 (1986).
22. McMillen, D. F. & Golden, D. M. *Rev. Phys. Chem.* **33**, 493–532 (1982).
23. Jones, W. D. & Feher, F. J. *J. Am. chem. Soc.* **106**, 1650–1663 (1984).
24. Martinho-Simoes, J. A. & Beauchamp, J. L. *Chem. Rev.* **90**, 629–688 (1990).

ACKNOWLEDGEMENTS. We thank S. Cohen for the X-ray structural analysis and M. Portnoy for discussions. The possible involvement of intermediate **10** in the reaction scheme in Fig. 4 was suggested by an anonymous referee. This work was supported by the Israel Science Foundation, administered by the Israel Academy of Sciences and Humanities.

Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment

James R. Ledwell*, Andrew J. Watson† & Clifford S. Law†

* Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

† Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, UK

THE distributions of heat, salt and trace substances in the ocean thermocline depend on mixing along and across surfaces of equal density (isopycnal and diapycnal mixing, respectively). Measurements of the invasion of anthropogenic tracers, such as bomb tritium and ³He (see, for example, refs 1 and 2), have indicated that isopycnal processes dominate diapycnal mixing, and turbulence measurements have suggested that diapycnal mixing is small^{3,4}, but it has not been possible to measure accurately the diapycnal diffusivity. Here we report such a measurement, obtained from the vertical dispersal of a patch of the inert compound SF₆ released in the open ocean. The diapycnal diffusivity, averaged over hundreds of kilometres and five months, was 0.11 ± 0.02 cm² s⁻¹, confirming previous estimates^{1–4}. Such a low diffusivity can support only a rather small diapycnal flux of nitrate into the euphotic zone; it justifies the neglect of diapycnal mixing in dynamic models of the thermocline^{25–27}, and implies that heat, salt and tracers must penetrate the thermocline mostly by transport along, rather than across, density surfaces.

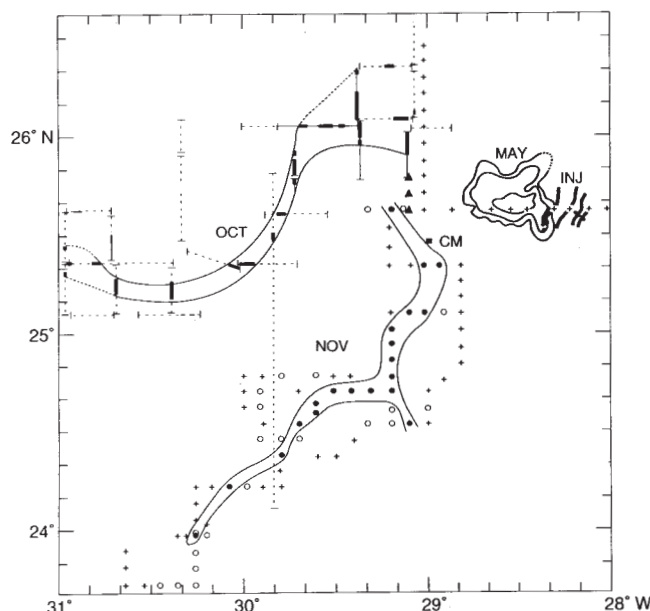


FIG. 1 Evolution of the lateral distribution of the tracer. The injection streaks are shown as short heavy lines near 26° N, 28° W. The contours just to the west show the patch later in May 1992. Heavy lines (further to the west) show tracks for the October survey, where the concentration C at the target surface was >500 fmol; light solid lines, C was between 100 and 500 fmol; dashed lines, C ~ 0. Solid triangles indicate bottle stations occupied at the end of the October cruise, with C > 300 fmol. Station symbols for the November survey are: plus signs, C < 30 fmol; open circles, C = 30–300 fmol; filled circles, C > 300 fmol. A fine curve has been drawn to envelop the high C regions for the two surveys. CM marks the location of the central mooring for the Subduction experiment.

Our experiment, which is a component of the World Ocean Circulation Experiment (WOCE), began with the release of 139 kg (950 mol) of sulphur hexafluoride (SF_6) on a target density surface about 310 m deep, at a location 1,200 km west of the Canary Islands. SF_6 is an inert and non-toxic compound⁵ which has proven to be conservative in pilot experiments^{6,7}. Yet it is detectable in amounts of less than 10^{-16} moles with a ship-board gas chromatograph equipped with an electron capture detector⁸. It is this sensitivity, and the oceanic background of less than 1 fM (10^{-15} M)⁹, which makes the large scale of the experiment possible.

The tracer was released from R/V *Oceanus* in a series of streaks between 5 and 13 May 1992 (INJ in Fig. 1). The injection package consisted of a pumping system that atomizes liquid SF_6 through a 50-micron sapphire orifice for efficient dissolution. It was towed at 0.5 m s^{-1} while being held automatically within a few metres of the target density surface.

Sampling of the tracer commenced from RRS *Charles Darwin*, 14–31 May 1992, by towing a vertical array of 19 samplers, spaced 2 to 6 m apart, through the patch at 0.5 m s^{-1} . Each sampler took in water over a period of ~ 4 h, the array yielding a vertical profile averaged over the track. An 18-chamber sampler was located at the centre of the array, each chamber integrating over ~ 350 m of the track. Virtually all of the tracer was found in a patch just west of the injection (MAY in Fig. 1), with the help of acoustically tracked floats released with the tracer¹⁰. The patch seemed to be made up of streaks of $\sim 1,500$ m width, with spacing between streak centres being of the order of 3,000 m; that is, the original injection streaks seemed well on their way towards filling in the patch area.

For the second survey, 26 September to 18 October from *Oceanus*, 23 samplers were spaced 4–6 m apart on the wire, the tow speed and duration were increased to 1.2 m s^{-1} and 7 h, respectively, and the central sampler had 50 chambers, each integrating laterally over 600 m. Ranges to the tracking floats were supplemented with previous estimates of the mean drift in the region^{11,12}, and with the drift of another nearby float provided by R. Davis. All of these sources indicated a mean drift of about 1 cm s^{-1} to the west or southwest, guiding the choice of starting point for the October survey near $24^\circ 07' \text{ N}$, $29^\circ 50' \text{ W}$ (Fig. 1).

The initial search encountered just one intense streak, about 8 km wide, at $25^\circ 29' \text{ N}$, $29^\circ 50' \text{ W}$. This streak was found to stretch at least 200 km, the west end branching into a number of thin, weak streaks, and the northeast end branching into a

number of wider, stronger limbs covering a region about 50 km across. Concentrations were of the order of 1,000 fM within the streak, and usually fell abruptly through 100 fM at the edges of the streak. Fifteen to twenty per cent of the tracer was found on this cruise.

It was found during the October cruise that the diapycnal distribution of the tracer could be sampled with conventional bottle casts made every 9 km. This finding implies a surprising amount of cross-streak homogeneity, and also that future surveys can cover an area more quickly. Indeed, bottle casts were used to continue the tracer survey from *Oceanus* in November, together with turbulence studies^{13,14}. The tracer was relocated near the east end of the streak found in October, and followed for 250 km to the southwest (NOV in Fig. 1). We estimate that 10–15% more of the tracer was found on this cruise, unless we resampled some of the same tracer encountered in October, which seems unlikely. If all of the tracer were distributed in a manner similar to that shown in Fig. 1, the total unfolded streak length would be $\sim 1,800$ km.

Figure 1 shows a clear separation between the mesoscale, represented by the radii of curvature of different parts of the streak, of the order of 100 km, and the finescale, represented by the r.m.s. cross-streak width, of the order of 3 km. According to a scale analysis¹⁵ and a numerical model¹⁶, this r.m.s. width implies an efficient process of dispersion at scales of 1 to 10 km. The unfolded length L of a tracer streak is thought to grow exponentially with time as the area occupied by the streak grows from the finescale to the mesoscale: $L = L_0 e^{\gamma t}$, where γ is of the order of the r.m.s. mesoscale strain rate. The value of γ corresponding to $L_0 = 40$ km from the first survey, and the $L = 1,800$ km estimated for the fall, is $3 \times 10^{-7} \text{ s}^{-1}$. The r.m.s. width σ_y of the streak is set in the models by a balance between lateral convergence associated with the strain and lateral eddy diffusion: $\sigma_y^2 = K_y/\gamma$, which gives an effective lateral eddy diffusivity K_y of the order of $3 \text{ m}^2 \text{ s}^{-1}$.

The streak width is not necessarily in steady state, and other processes such as cross-streak shear in the along-streak direction could be elongating the streak, as has been pointed out by E. Kunze (personal communication). But any scheme requires K_y to be of the order of $1 \text{ m}^2 \text{ s}^{-1}$ to give the 40-fold increase in the area actually stained by the tracer and the concomitant decrease in maximum concentration over the 5-month period.

This value of K_y is at least an order of magnitude larger than the small scale lateral eddy diffusivity expected from the interaction between vertical shear of the internal waves and vertical mixing¹⁷. Weaker but more persistent vertical shears, associated with the mesoscale eddy field or small scale vortices¹⁸, could be responsible for the observed lateral dispersion, however¹⁷.

Each survey shown in Fig. 1 produced 20 to 30 vertical tracer profiles, which were averaged as a function of height h above the target density surface (Fig. 2). An estimate of the diapycnal eddy diffusivity K_z can be obtained from the increase of the second moment M_2 of these mean profiles with time t according to: $2 K_z = dM_2/dt$.

Values for M_2 , estimated as the second moment of gaussian functions fitted to the mean profiles, are shown in Fig. 3 for two cases. The upper point for each survey is M_2 calculated relative to the target density surface; the lower point is M_2 relative to the centre of mass, which is shifted downwards by 0.8 m, 2.7 m and 5.1 m for the May, October and November surveys, respectively. Use of the lower points would be appropriate if these shifts were artefacts, whereas the upper ones would be appropriate if the shifts are the result of an oceanographic phenomenon, such as shear distortion. The difference in the value of K_z inferred from the best-fit slope for the two cases is only 7%. The middle ground results in an estimate of $K_z = 0.11 \pm 0.02 \text{ cm}^2 \text{ s}^{-1}$. This result agrees with preliminary estimates of the eddy diffusivity of density and temperature from measurements of turbulent dissipation rates during a broad site survey in April 1992 (ref. 19).

There are adjustments to be made to this estimate of K_z owing

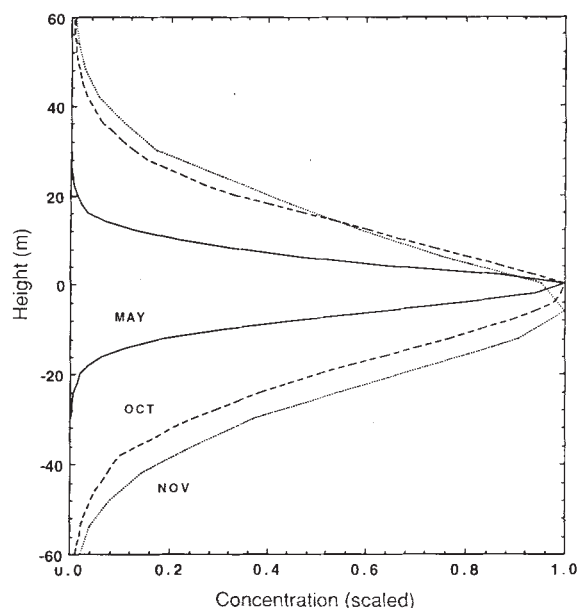


FIG. 2 Evolution of the vertical distribution of the tracer. The mean profiles have been scaled so that the widths can be compared.

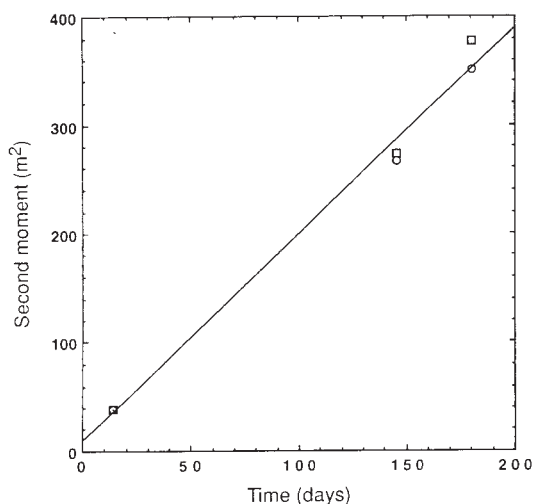


FIG. 3 Growth of the second moment of the vertical tracer distribution. Squares are for raw M_2 , circles are for the centre of mass shifted to $h=0$. The line is for $K_z = 0.11 \text{ cm}^2 \text{ s}^{-1}$.

to the variation of density at fixed h , and to changes in mean stratification between May and November. Preliminary estimates indicate that these adjustments nearly cancel one another and contribute most of the uncertainty in K_z .

This estimate of K_z applies to the part of the ocean sampled by the tracer we found, namely a volume of space-time of the order of $10^{17} \text{ m}^3 \text{ s}$. One cannot prove that the unsampled part of the patch has the same vertical distribution. But the evidence from the site survey is that the turbulent dissipation rates were uniform over the region¹⁹. Measures of temporal variability in the dynamics driving the mixing will be available from current metres on the central mooring²⁰ of the Subduction Experiment, another component of WOCE, (Fig. 1), and from a specially designed float²¹.

These data promise extrapolation of our measurement of K_z to a large fraction of the upper 1,000 m of the global ocean in stratified waters below the region mixed by cooling each winter. This stratified layer is important as a reservoir of nutrients for the euphotic zone, and as a capacitor for perturbations of atmospheric temperature and carbon dioxide associated with climate change.

If a value for K_z of $0.11 \text{ cm}^2 \text{ s}^{-1}$ were to prove typical of the upper ocean, then diapycnal mixing would not be the dominant transport mechanism within this layer for many processes. For example, the flux of nitrate through the upper layer to support 'new' biological productivity in the euphotic zone would be of the order of $0.01 \text{ M m}^{-2} \text{ yr}^{-1}$. Although this flux is about 50% of that required by estimates of productivity for the subtropical oceans based on incubations²², it is only 2% of the flux required by *in situ* oxygen and ^3He budgets²³. Thus, other nitrate transport mechanisms may be operating (see, for example ref. 24).

Also, if K_z is so small, the dynamics of the upper ocean can be modelled by neglecting diapycnal processes in the stratified interior, as it often has been^{25–27}. Furthermore, the response of this layer to a perturbation of atmospheric temperature or carbon dioxide would not be governed by diapycnal mixing in the interior. Rather, transport of nitrate, heat, salt and trace constituents would be dominated by advection along density surfaces, coupled with mixing at the surface and other boundaries.

A final sampling of the tracer patch in the spring of 1993 will provide a measurement of K_z for a winter period, and further information about lateral processes. Future experiments, especially in the deeper ocean and near boundaries, could clarify where mixing is occurring and its strength.

Note added in proof: Virtually all the tracer was found in the spring, in a 600,000-km² patch rich with streaks. Preliminary

estimates are that K_z for the entire year was $0.15 \text{ cm}^2 \text{ s}^{-1}$, suggesting somewhat more vigorous mixing in the winter. □

Received 10 February; accepted 25 June 1993.

1. Rooth, C. G. & Ostlund, H. G. *Deep-Sea Res.* **19**, 481–492 (1972).
2. Jenkins, W. J. *J. Mar. Res.* **38**, 533–569 (1980).
3. Moun, J. N. & Osborn, T. R. *J. phys. Oceanogr.* **16**, 1250–1259 (1986).
4. Gregg, M. C. *J. geophys. Res.* **94**, 9686–9698 (1989).
5. Lester, D. & Greenberg, L. A. *Arch. Ind. Hyg. Occup. Med.* **2**, 348–349 (1950).
6. Ledwell, J. R. & Watson, A. J. *J. geophys. Res.* **96**, 8695–8718 (1991).
7. Watson, A. J., Ledwell, J. R. & Sutherland, S. C. *J. geophys. Res.* **96**, 8719–8725 (1991).
8. Wanninkhof, R., Ledwell, J. R. & Watson, A. J. *J. geophys. Res.* **96**, 8733–8740 (1991).
9. Watson, A. J. & Liddicoat, M. I. *Atmos. Environ.* **19**, 1477–1484 (1985).
10. Price, J., McKee, T., Valdes, J., Richardson, P. & Armi, L. Report 86–31, Woods Hole Oceanographic Institution (Woods Hole, MA, 1986).
11. Armi, L. & Stommel, H. *J. phys. Oceanogr.* **13**, 828–857 (1983).
12. Thiele, G. et al. *J. phys. Oceanogr.* **16**, 814–826 (1986).
13. Oakey, N. S. *IEEE J. Ocean. Engng* **13**, 124–128 (1988).
14. Duda, T. F., Cox, C. S. & Deaton, T. K. *J. Atmos. Ocean Tech.* **5**, 16–33 (1988).
15. Garrett, C. *Dyn. Atmos. Oceans* **7**, 265–277 (1983).
16. Haidvogel, D. B. & Keffer, T. *Dynam. Atmos. Oceans* **8**, 1–40 (1984).
17. Young, W. R., Rhines, P. B. & Garrett, C. J. R. *J. phys. Oceanogr.* **12**, 515–527 (1982).
18. Muller, P., Lien, R.-C. & Williams, R. J. *J. phys. Oceanogr.* **18**, 401–416 (1988).
19. Montgomery, E., Schmitt, R. W., Toole, J. M. & Polzin, K. L. *EOS* **73**, 321 (1992).
20. Hosom, D. S., Weller, R. A., Prada, K. E. & Trask, R. P. *Proc. Mar. Tech. Soc.* **1**, 206–210 (1991).
21. Williams, A. J., Converse, C. H. & Nicholson, J. *Am. Soc. mech. Engng, OED* **12**, 25–29 (1987).
22. Eppley, R. W. & Peterson, B. J. *Nature* **282**, 677–680 (1979).
23. Jenkins, W. J. *Nature* **331**, 521–523 (1988).
24. Vilareal, T. A., Altabet, M. A. & Culver-Rymsa, K. *Nature* **363**, 709–712 (1993).
25. Welander, P. *Tellus* **11**, 309–318 (1959).
26. Luyten, J. R., Pedlosky, J. & Stommel, H. *J. phys. Oceanogr.* **13**, 292–309 (1983).
27. Huang, R. X. *Rev. Geophys. suppl.* 590–609 (1991).

ACKNOWLEDGEMENTS. The tracer experiment is an outgrowth of suggestions by W. S. Broecker, J. Shepherd and J. Lovelock. We thank S. Becker, D. Clochetto, E. Clark, K. Doherty, T. Donoghue, K. Fairhurst, C. Fernandez, B. Guest, S. Kery, M. Kias, C. Kinkade, M. Liddicoat, K. Lubcke, C. Marquette, A. Martin, P. Nightingale, N. Oakey, R. Oxburgh, H. Rochat, B. Ruddick, S. Sutherland, K. Tedesco, J. White, RVS in the UK, the crews of *Darwin* and *Oceanus*, our home laboratories for their help, and E. Kunze for comments on the manuscript. The experiment has been supported by the US NSF and the UK Natural Environment Research Council.

A $\delta^{13}\text{C}$ record of late Quaternary climate change from tropical peats in southern India

R. Sukumar*, R. Ramesh†, R. K. Pant† & G. Rajagopalan‡

* Centre for Ecological Sciences, Indian Institute of Science, Bangalore–560 012, India

† Physical Research Laboratory, Ahmedabad–380 009, India

‡ Birbal Sahni Institute of Palaeobotany, Lucknow–226 007, India

STABLE-ISOTOPE ratios of carbon in soils or lake sediments^{1–3} and of oxygen and hydrogen in peats^{4,5} have been found to reflect past moisture variations and hence to provide valuable palaeoclimate records. Previous applications of the technique to peat have been restricted to temperate regions, largely because tropical climate variations are less pronounced, making them harder to resolve. Here we present a $\delta^{13}\text{C}$ record spanning the past 20 kyr from peats in the Nilgiri hills, southern India. Because the site is at high altitude (>2,000 m above sea level), it is possible to resolve a clear climate signal. We observe the key climate shifts that are already known to have occurred during the last glacial maximum (18 kyr ago) and the subsequent deglaciation. In addition, we observe an arid phase from 6 to 3.5 kyr ago, and a short, wet phase about 600 years ago. The latter appears to correspond to the Mediaeval Warm Period, which previously was believed to be confined to Europe and North America^{6,7}. Our results therefore suggest that this event may have extended over the entire Northern Hemisphere.

The use of stable carbon isotope ratios as palaeoclimatic indicators is based on the different ecological requirements of the C3 and C4 plant types that have widely differing $^{13}\text{C}/^{12}\text{C}$ ratios.

C3 and C4 plants, separated on the basis of their photosynthetic pathways of carbon fixation, typically have $\delta^{13}\text{C}$ values in the range of -26‰ to -28‰ and -11‰ to -13‰ respectively⁸ ($\delta^{13}\text{C}$ is defined in the legend of Fig. 2). They also show different ecological preferences, with C4 plants (mainly tropical grasses) favouring conditions of aridity and low soil moisture and C3 plants (most dicotyledonous plants and temperate grasses) dominating areas of higher precipitation and higher soil moisture^{9,10}. Although the precise ecological separation of these two plant types in the tropics is not clearly understood, the C4 grasses are known to predominate up to an altitude of 2,500 m above mean sea level (m.s.l.) under conditions of low soil moisture, whereas C3 grasses predominate at greater altitudes, where soil moisture is greater⁹. The balance between C3 and C4 grasses depends on soil moisture and not altitude itself. For this reason, $\delta^{13}\text{C}$ values of soil organic matter, calcium carbonate or lake sediments (as they relate to C3 and C4 vegetation) have been used as palaeoclimatic indicators^{1,3}.

The current vegetation of the high altitude (1,800–2,500 m above m.s.l.) regions of the Nilgiri hills (11° – $11^{\circ} 30'$ N, 76° – 77° E) in the Western Ghats of southern India is a mixture of stunted evergreen forest and grassland^{11,12}. The undulating hills feature an underlying geology of Archaean rocks, chiefly gneisses, charnockites and schists. The western region receives rain mainly from the south-west (summer) monsoon and the eastern region from the north-east (winter) monsoon, with a gradation over the plateau. Present-day rainfall over the plateau varies from about 3,000 mm per year in the west to only 1,200 mm per year in the more sheltered central basins¹³. Mean monthly temperatures typically vary from a maximum of 20°C in May to 7°C in January. Frost is however common during the month of January and limits the spread of forest plants^{11,12}. A number of climate models have shown a positive correlation between temperature and the strength of the Asian summer monsoon^{14,16}. In general, a higher mean temperature would reduce the incidence of frost, increase summer precipitation and promote the spread of forest (C3 type) and C3 grasses in moist soils of grasslands. Conversely, a cooler mean temperature would favour the spread of grassland (C4 grasses in low soil moistures at this altitude). Thus, a strengthening or weakening of the monsoon could be expected to be reflected in a swing towards C3 or C4 vegetation respectively.

The valleys feature peat deposits whose formation can be dated back to at least 40 kyr BP (before present) (refs 17–20, and our unpublished results). The peats derived from the C3 and C4 vegetation of the surrounding area can thus be expected to preserve the vegetational and climatic record of the region, as diagenetic alteration usually does not occur⁷.

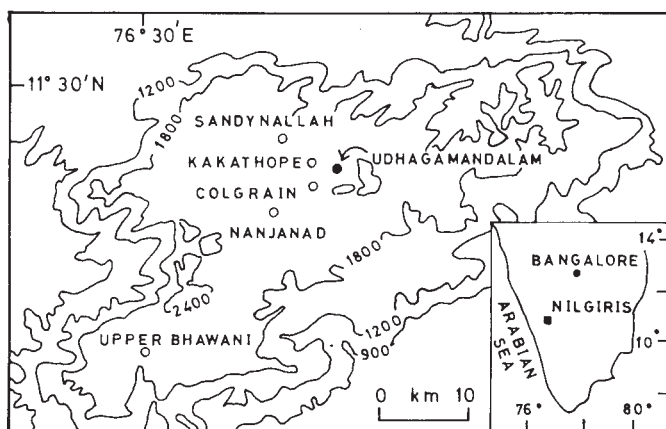


FIG. 1 Map showing the locations of the basins (open circles) in the Nilgiri hills from where peat samples were collected. Contours are given in metres above mean sea level.

TABLE 1 Ages for peat samples from the Nilgiri basins

Basin	Depth (cm)	Lab. sample ref. no.	^{14}C age ($\pm$ s.d.) (yr BP)	Recalibrated age (yr BP)
Upper Bhawani	0–5	BS-76	290 ± 100	311
	20–40	BS-75	$1,980 \pm 100$	1,935
	130–150	BS-52	$5,860 \pm 115$	6,700
	200–230	BS-53	$19,100 \pm 300$	22,300
Kakathope	70–150	BS-196	$5,600 \pm 120$	6,370
	120–150	BS-187	$7,760 \pm 140$	8,540
	170–200	BS-197	$10,620 \pm 180$	12,280
	220–250	BS-188	$14,480 \pm 240$	17,320
Nanjanad	20–50	BS-106	$4,130 \pm 100$	4,700
	120–150	BS-122	$19,890 \pm 380$	23,000
Colgrain I	20–50	BS-167	$7,580 \pm 130$	8,380
	70–100	BS-168	$17,140 \pm 290$	19,035
Colgrain II	220–250	BS-23	$14,920 \pm 960$	18,000
Sandynallah	5–10	BS-874	Modern	—
	150–155	BS-873	$2,230 \pm 130$	2,315
	160–165	BS-968	$2,550 \pm 90$	2,739
	170–175	BS-871	$3,040 \pm 80$	3,263
	180–185	BS-967	$4,050 \pm 110$	4,500
	190–195	BS-872	$4,680 \pm 120$	5,390
	233–240	BS-869	$8,680 \pm 140$	9,500

Visible plant debris were removed from the samples before processing for radiocarbon counts (ref. 17). Several samples from the upper layers of the Sandynallah basin, after removal of visible modern rootlets, did not have sufficient carbon for accurate dating; these have been dated by calculating rates of peat deposition from least-squares regression. The best fit was given by $\log_e y = 0.0165x + 5.207$ ($R^2 = 0.986$), where y is the age in yr and x the depth in cm. Peats between 195 and 227 cm at Sandynallah were lost during collection. One sample at 227–233 cm depth has been dated from the regression. All ^{14}C ages are based on a half-life of $5,730 \pm 40$ yr. The recalibrated ages are mean ages based largely on dendrochronological data (ref. 21) or U/Th ages of corals (ref. 22) and are approximate for ages greater than 4,000 years. The reference year for BP is AD 1950.

We collected peat samples, each of 5 cm depth, at 10 cm intervals from a 2.4-m-deep pit dug in the Sandynallah basin. Additional samples were obtained from the collections at the Birbal Sahni Institute of Palaeobotany (BSIP)^{17–19}, which had been made by using a corer (for pollen studies) at Kakathope, Nanjanad, Colgrain and Upper Bhawani (Fig. 1). The general stratigraphy of these basins has been described elsewhere^{17–20}. Radiocarbon dating of the samples was carried out at BSIP by methods described in ref. 15. Table 1 shows these ^{14}C ages. In order to correct the radiocarbon ages for past variations in atmospheric ^{14}C , we also indicate recalibrated ages based on tree-ring dates²¹ or U/Th dates of corals²². In the discussion here, however, we refer to conventional ^{14}C ages in order to facilitate comparison with past studies which report only radiocarbon ages.

For determining $^{13}\text{C}/^{12}\text{C}$ ratios, samples were combusted in sealed quartz tubes at 800°C , and the resulting CO_2 analysed (with a resolution of $\pm 0.1\text{‰}$ (l.s.d.) in a Micromass 602D mass spectrometer at the Physical Research Laboratory. All $\delta^{13}\text{C}$ values employ the PDB standard. The fine and coarse fractions (usually vegetal debris) of a number of peat samples were separately analysed to check for consistency: the two values matched closely in all cases. We therefore report the $\delta^{13}\text{C}$ values of the fine fraction which is more representative of the carbon contribution of the prevalent vegetation.

Values of $\delta^{13}\text{C}$ in present-day dominant vegetation in the Nilgiri basins average -10.9‰ (± 1.22 , number of samples $n = 5$) for C4 grasses, -27.8‰ (± 2.27 , $n = 7$) for C3 grasses and other herbs, and -28.7‰ (± 1.32 , $n = 8$) for C3 trees and shrubs. Modern (sub-surface) peats average -18.1‰ (± 1.17 , $n = 7$).

Figure 2 shows the $\delta^{13}\text{C}$ values of the peats collected from the Nilgiri basins. The $\delta^{13}\text{C}$ values have fluctuated in the past from -12.8‰ to -24.2‰ indicating a dominance of C4 or C3 vegetation respectively. The predominantly C4 signatures during 20–16 kyr BP (Fig. 2a) clearly indicate a very arid phase during the last glacial maximum (LGM, believed to be at 18 kyr BP). Evidence from other palaeoclimatic data (including lacustrine pollen deposits in northwestern India²³, oxygen isotope ratios of oceanic sediments and foraminifera^{24,27}, and climatic simulation models^{14,16}) also points to a period of weak (compared to the

present) Asian south-west (summer) monsoon during the LGM. This period is accompanied by cooler global temperatures. The north-east (winter) monsoon was perhaps stronger during this period^{26,27} but the net precipitation from south-west and north-east monsoons over the Western Ghats of peninsular India could still have been relatively lower then.

From about 16 kyr BP there is a tendency towards a more negative $\delta^{13}\text{C}$ signature in the peat, which indicates C3 vegetation dominance, attaining a peak of -24.2‰ at 10.6 kyr BP at Kakathope (Fig. 2a). This matches the strengthening of the summer monsoon reported for this period from the following sources; pollen records in lakes of northwestern India^{23,28}, fluvial deposits in central India²⁹, oxygen isotope and pollen records from cores in the Arabian Sea^{24,25}, and simulation models based on increased solar radiation (such as at 9 kyr BP) caused by the earth's orbital variation^{16,30}. This evidence indicates that the summer monsoon reached a peak at ~ 11 kyr BP at $10\text{--}15^\circ\text{N}$ latitude, and ~ 10 kyr BP at more northern latitudes in the sub-continent. The extension of the monsoon over the Indian sub-continent seems to have been a gradual process, following the northward progression of the intertropical convergence zone during the deglaciation²⁵.

The early Holocene (10–6 kyr BP) is again marked by a shift towards C4 vegetation, indicating a progressively more arid climate. This trend is clearly captured in the closely-sampled Sandynallah basin (Fig. 2b), where the $\delta^{13}\text{C}$ value of -20.2‰

at 8.7 kyr BP shows a positive excursion to between -14.7‰ and -15‰ during 4.7–3.0 kyr BP. At Nanjanad, a $\delta^{13}\text{C}$ value of -14.3‰ at 4 kyr BP (Fig. 2a) was found. The weakening of the summer monsoon over the Indian sub-continent has been variously dated at 4.5–2 kyr BP in northern India at latitude 28°N ^{23,25,28}. A sharp increase in grass pollen, indicating aridity from 3.5 kyr BP onwards, has also been reported from estuarine sediments off the western coast of peninsular India at latitude 15°N (ref. 31). Our data indicate that a weaker monsoon may have established as early as 6 kyr BP in peninsular India at latitude 11°N .

The Sandynallah basin also shows a sudden shift in $\delta^{13}\text{C}$ values (beginning ~ 0.7 kyr BP) to a C3 vegetation (extreme $\delta^{13}\text{C}$ of -22.6‰ at 0.6 kyr BP) for a duration of about 200 yr (Fig. 2b). After correcting these with dendrochronological ages²¹, this period (calendar age AD 1200–1400) roughly corresponds to the Mediaeval Warm Period, dated at AD 1100–1300, characterized by warmer temperatures in Europe and North America^{6,7}. Our record seems to be the first evidence of such an event from southern Asia, although its relationship to the better known European event is unclear at this stage.

The broad trends in monsoonal evolution thus seem to be captured in the Nilgiri peats. There could however be pitfalls in the interpretation of peat records. The stratigraphy of two basins on the western fringe of the plateau (that receives torrential rains from the summer monsoon) showed anomalies (through ^{14}C date inversions and contamination with modern carbon) that could have been caused by landslides triggered by a high-energy event (A. K. Singhvi, personal communication). More sheltered basins, such as those investigated in this work, would be better sampling sites. It could also be important to sample peat deposits at altitudes lower than 2,500 m. At higher altitudes a possible predominance of C3 over C4 grasses, due to persistently higher soil moisture, may render the carbon isotope signature ineffective. Natural lag in vegetational response to climatic change should also be taken into account.

Tropical climates are generally more difficult to reconstruct than temperate climates, for a number of reasons. Changes in ocean temperatures in the tropics are very small. Seasonal contrast of climate is also marginal at lower altitudes. For these reasons an elevated region with higher diurnal and seasonal contrast can be expected to give a better record of climate change.

We have demonstrated the potential of using tropical peats as palaeoclimatic indicators. We believe that further extensive studies, using a combination of pollen and stable isotopes of C, O and H may be very useful tools in past reconstruction under the International Geosphere-Biosphere Programme (IGBP). This may allow us to extend the geographic coverage of proxy climatic sets. □

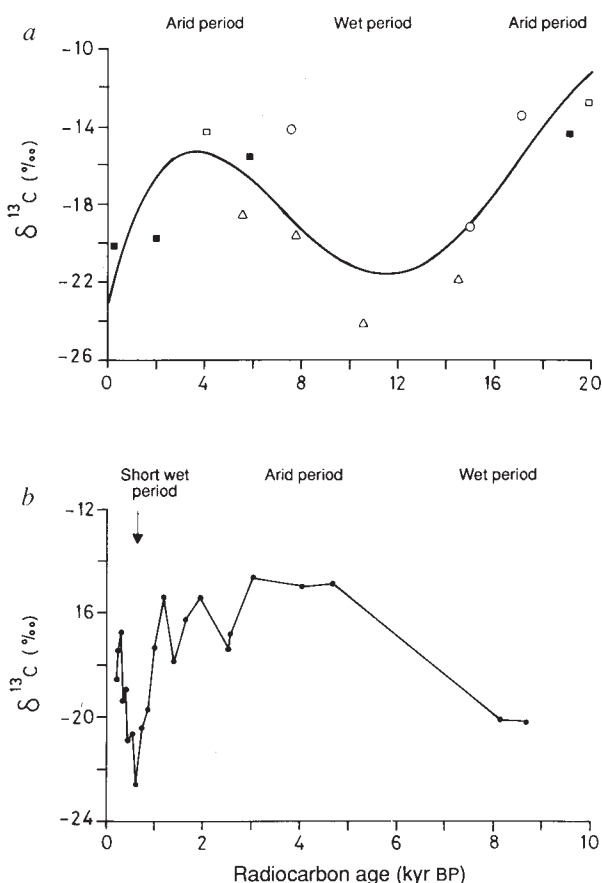


FIG. 2 $\delta^{13}\text{C}$ (per mil), relative to PDB standard, versus radiocarbon date (kyr BP) of peats from the Nilgiri basins. a, Samples collected by BSIP (refs 17–19) from Nanjanad (open square), Colgrain (open circle), Kakathope (open triangle) and Upper Bhawani (closed square). A fourth-degree polynomial function has been fitted to the data to depict the broad trends in $\delta^{13}\text{C}$ since 20 kyr BP. b, Samples collected from the Sandynallah basin with ages 9 kyr BP to present. ($\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}] - 1$).

Received 27 January; accepted 25 June 1993.

- Krishnamurthy, R. V., Bhattacharya, S. K. & Kusumgar, S. *Nature* **341**, 150–152 (1986).
- Cerling, T. E., Quade, J., Wang, Y. & Bowman, J. R. *Nature* **341**, 138–139 (1989).
- Ambrose, S. H. & Sikes, N. E. *Science* **253**, 1402–1405 (1991).
- Brenninkmeijer, C. A. M., van Geel, B. & Mook, W. G. *Earth Planet. Sci. Lett.* **61**, 283–290 (1982).
- Dupont, L. M. & Mook, W. G. *Chem. Geol. (Isotope Geosci. Sect.)* **66**, 323–333 (1987).
- Gribbin, J. & Lamb, H. H. in *Climatic Change* (ed. Gribbin, J.) 68–82 (Cambridge Univ. Press, 1978).
- Crowley, T. J. & North, G. *Palaeoclimatology* (Oxford Univ. Press, New York, 1991).
- Smith, B. N. & Epstein, S. *Plant Physiol.* **47**, 380–384 (1971).
- Tieszen, L. L., Senyimba, M. M., Imbamba, S. K. & Troughton, J. H. *Oecologia* (Berlin) **37**, 337–350 (1979).
- Korner, Ch., Farquhar, G. D. & Roksandic, Z. *Oecologia* (Berlin) **74**, 623–632 (1988).
- Blasco, F. *Montagnes du Sud de l'Inde* (Institut Français, Pondicherry, 1971).
- Meher-Homji, V. M. *Indian geogr. J.* **59**, 205–213 (1984).
- Von Lengerke, H. J. *The Nilgiri—Climate and Weather of a Mountain Area in South India* (Steiner, Weisbaden, 1977).
- Gates, W. L. *Science* **191**, 1138–1144 (1976).
- Manabe, S. & Hahn, D. G. *J. geophys. Res.* **82**, 3889–3911 (1977).
- Prell, W. L. & Kutzbach, J. E. *J. geophys. Res.* **92**, 8411–8425 (1987).
- Rajagopalan, G., Vishnu-Mittre & Sekar, B. *Radiocarbon* **20**, 398–404 (1978).
- Rajagopalan, G., Vishnu-Mittre & Sekar, B. *Radiocarbon* **22**, 54–60 (1980).
- Rajagopalan, G., Vishnu-Mittre, Sekar, B. & Mandal, T. K. *Radiocarbon* **24**, 45–53 (1982).
- Vasanthy, G. *Rev. Palaeobot. Palynol.* **55**, 175–192 (1988).
- Stuiver, M. & Kra, R. S. (eds) *Proc. Twelfth Int. Radiocarbon Conf.* (Trondheim, Norway: published as *Radiocarbon* **28**, 805–1030 (1986).

22. Bard, E., Hamelin, B., Fairbanks, R. G. & Zindler, A. *Nature* **345**, 405–410 (1990).
 23. Singh, G., Wasson, R. J. & Agrawal, D. P. *Rev. Palaeobot. Palynol.* **64**, 351–358 (1990).
 24. Van Campo, E., Duplessy, J. C. & Rossignol-Strick, M. *Nature* **296**, 56–59 (1982).
 25. Van Campo, E. *Quat. Res.* **26**, 376–388 (1988).
 26. Duplessy, J. C. *Nature* **295**, 494–498 (1982).
 27. Sarkar, A., Ramesh, R., Bhattacharya, S. K. & Rajagopalan, G. *Nature* **343**, 549–551 (1990).
 28. Swain, A. M., Kutzbach, J. E. & Hastenrath, S. *Quat. Res.* **19**, 1–17 (1983).
 29. Williams, M. A. J. & Clarke, M. F. *Nature* **308**, 633–635 (1984).
 30. Kutzbach, J. E. *Science* **214**, 59–61 (1981).
 31. Caratini, C., Fontugne, M., Pascal, J. P., Tissot, C. & Bentaleb, I. *Curr. Sci.* **61**, 669–672 (1991).

ACKNOWLEDGEMENTS. This work was partially supported by the Ministry of Environment and Forests, Government of India. We thank the Forest Department of Tamilnadu and the Sheep Breeding Centre, Sandynallah, for permission to collect peat samples. We also thank J. Srinivasan, N. V. Joshi and A. K. Singhvi for suggestions and J. P. Padia, G. Rajagopalan and J. A. Santosh for assistance.

Dependence of ridge-axis morphology on magma supply and spreading rate

Jason Phipps Morgan* & Y. John Chen†

* Institute of Geophysics and Planetary Physics, UCSD, La Jolla, California, USA

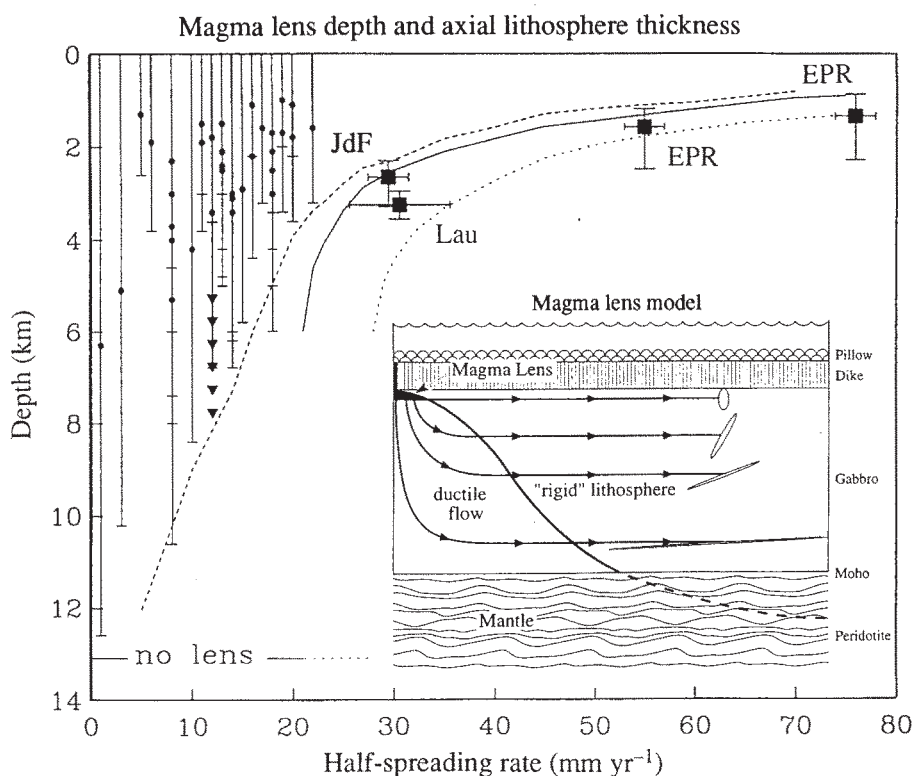
† College of Oceanography, Oregon State University, Corvallis, Oregon, USA

WHY do some mid-ocean-ridge axes have as their topographic expression a median valley 1–2 km deep and 15–30 km wide, whereas others have an axial high 100–200 m high and 1–2 km wide? In general, median valleys exist at the axes of ridges spreading at half-rates less than ~ 20 – 25 mm yr^{-1} , whereas axial highs are found at faster-spreading ridges^{1,2}. A recent crustal genesis model³, incorporating hydrothermal cooling and crustal accretion by means of a magma lens, agrees well with observations of the spreading-rate dependence of axial morphology^{1,2}, the depth of the axial magma chamber^{3,4} and the thickness of the ridge-axis

lithosphere^{5,6}. But there are several notable exceptions to the correlation between spreading rate and axial morphology, such as the Reykjanes Ridge, which spreads at 10 mm yr^{-1} but has an axial high. Here we show that by extending the model of ref. 3 to include variations in crustal thickness, we can explain the observed dependence of axial morphology on crustal thickness and spreading rate. Our results suggest that the ultimate control on axial morphology is the thermal structure at the ridge axis, which is a function of both spreading rate and magma supply.

Recently, the original 'magma chamber' and crustal flow model of Sleep⁷ has been re-examined. Magma injection into a small magma lens at the base of the sheeted-dike complex⁸ (see inset to Fig. 1) has the potential to explain both the gabbro layering patterns seen in the ophiolite record and the small magma lens that seismic methods image at active fast-spreading ridges^{3,9,10}. Thermal predictions and geophysical observations are reconciled by including the effects of hydrothermal circulation in the heat budget^{3,10}. The spreading-rate dependence of axial topography is explained as follows: (1) stretching of strong axial lithosphere is responsible for the origin of a median valley along a spreading centre, and (2) the thickness of this strong lithosphere is controlled by a delicate balance between heat input

FIG. 1 Depth to the top of the magma lens as a function of spreading rate, all other parameters being held constant. Curves show results from a suite of numerical experiments with $Nu=8$ and $Nu=12$ and $T_{\text{cutoff}}=600^\circ\text{C}$. For these hydrothermal cooling parameters a steady-state magma lens can only exist within the crust at half-spreading rates greater than about 20 (or 30) mm yr^{-1} . A well-developed shallow magma lens only exists for spreading rates greater than 30 mm yr^{-1} . Filled squares (and associated uncertainties from ref. 4) show multichannel seismic observations of the depth of the magma lens along intermediate- and fast-spreading ridges^{3,4} (EPR is East Pacific Rise, Lau is Lau back-arc-basin ridge and JdF is Juan de Fuca ridge). There is good agreement between model predictions of the depth-dependence of the magma lens with spreading rate, and multichannel seismic observations. Depths to the 750°C isotherm are also shown for a suite of numerical experiments with $Nu=8$ and $T_{\text{cutoff}}=600^\circ\text{C}$. The model 750°C isotherm correlates well with the spreading-rate dependence inferred from teleseismic axial earthquake centroid depths²⁸ (shown by small dots and associated bars representing inferred total rupture depth). Triangles are microearthquake focal depths beneath the axis of the northern Mid-Atlantic Ridge^{5,6}. There is good agreement between model isotherms and the spreading-rate dependence of the depth of the seismically observed brittle-ductile transition. The inset shows a schematic of the model of crustal genesis used for these results³; showing the flowlines and accumulated strain predicted by lower crustal flow away from a shallow



injection site at the sheeted dike/gabbro interface. Strain ellipses parallel the direction of strain-induced gabbro layering. Strain is most intense in the lowermost part of the gabbro section. See text for further discussion of the model.

by injection of magma into the crust and removal by hydrothermal circulation, so that the presence or absence of axial topography directly reflects the axial thermal structure along a ridge.

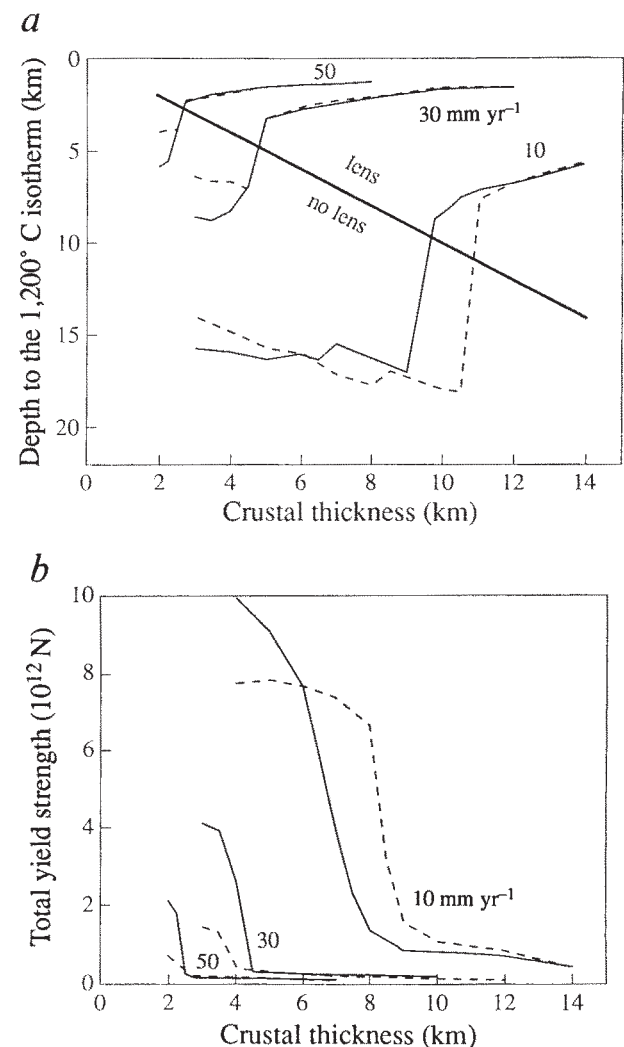
The thermal structure at a spreading centre is predominantly influenced by two factors: (1) the rate and geometry of magma supply, and (2) the efficiency of hydrothermal circulation in removing heat through rocks that are cool enough to permit cracking and hydrothermal heat transport. We treat magma ascent through hot host rock as a buoyant ascent, which is halted only where melt reaches a freezing surface¹¹ beneath which it ponds and forms a small melt lens. The time-averaged thermal effect of sheeted-dike injection and pillow flows is modelled by a 250-m-wide axial 'dike' that is emplaced at the plate spreading rate. The rest of the oceanic crust is injected into a steady-state magma lens directly beneath the 'solidus' freezing horizon (here taken to be 1,200 °C) at a rate sufficient to supply all crust which is not emplaced through diking/extrusion above the magma lens. We take the ~1-km-wide, ~250-m-thick prismatic shape found in recent East Pacific Rise (EPR) seismic studies¹²⁻¹⁴ as a kinematic constraint on the shape of this magma lens. All the latent heat release associated with the freezing of the lower crust is assumed to occur in the lens, because later freezing of the small (<3–5%) intracumulate melt fraction that remains below the melt lens¹⁵ releases less than 3–5% of the total heat released. Note that a crustal magma lens will cease to exist in this model if the steady-state thermal structure places this solidus isotherm beneath the crust. (In this case, a persistent magma body may still exist at the magma solidus depth within the mantle, but

freezing within this body does not directly contribute to crustal emplacement though it may shape magma chemistry. Where hydrothermal cooling occurs, the ordinary thermal conductivity is enhanced by a factor Nu , the Nusselt number, which is the ratio of hydrothermal heat transport within a permeable layer to heat transport by heat conduction alone. Rock at temperatures greater than the hydrothermal temperature limit T_{cutoff} ($=600$ °C) is assumed to be impermeable to hydrothermal flow. We also assume, unless otherwise stated, that the maximum depth of hydrothermal penetration Z_{cutoff} is pressure-limited, even when $T < T_{\text{cutoff}}$, to depths at which cracks can stay open; a limit that occurs at roughly Moho-depths. (This depth-limit is supported by oxygen isotope depletion studies of palaeo-hydrothermal circulation in ophiolites¹⁶⁻¹⁷.) The hydrothermal convection Nu is taken in the range 8–12 to match the observed magma lens depth of ~1.5 km along the northern EPR³ (compare Fig. 1). Note that the depth of the EPR magma lens does not appear to lie at the magma's neutral buoyancy level within the crust¹⁸.

To explore the effect of variations in magma supply at a given spreading rate we have done a suite of calculations at half-spreading rates of 10, 30 and 50 mm yr⁻¹ for a range of potential crustal thicknesses. Figure 2 shows these results for $T_{\text{cutoff}} = 600$ °C, $Nu = 8$, and two assumptions about the maximal depth penetration of hydrothermal circulation: (1) hydrothermal penetration is limited to ~6 km independent of crustal thickness, and (2) hydrothermal penetration always has the potential to just reach the Moho, where cracks become rapidly filled with precipitates.

The axial strength of the lithosphere shown in Fig. 2b is the

FIG. 2 Magma lens depth and axial yield strength plotted as a function of crustal thickness at spreading rates of 10, 30, and 50 mm yr⁻¹. **a** shows the depth of the 1,200 °C isotherm plotted as a function of crustal thickness for $Nu = 8$, $T_{\text{cutoff}} = 600$ °C, and for two different hydrothermal penetration limits, either $Z_{\text{cutoff}} = 6$ km (solid line) or $Z_{\text{cutoff}} =$ the crustal thickness (dashed line). The region above the lens/no lens transition is where the 1,200 °C isotherm will lie within the crust. Note the sharp change in the depth of this isotherm for small variations in crustal thickness around the crustal melt lens–no crustal lens transition. **b** The corresponding integrated axial yield strength plotted as a function of crustal thickness. The axial yield strength also changes dramatically near this threshold magma supply rate.



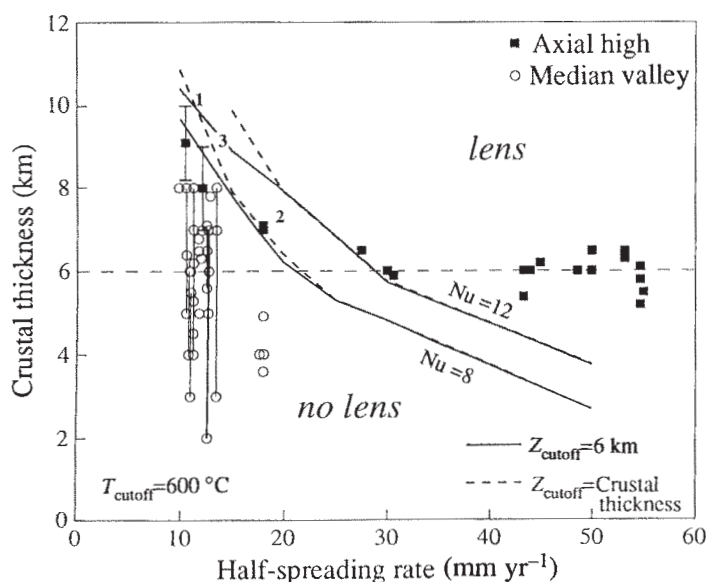


FIG. 3 The minimum crustal thickness where a steady-state magma lens will exist within the crust is plotted against spreading rate. Axial morphology and corresponding axial crustal thickness determined from seismic measurements are also shown as a function of spreading rate. Axial high morphology is shown by filled squares, median valley morphology is shown by open circles. Circles connected by solid lines show the range of crustal thickness variations seen along several fracture zone profiles. All data are a near-ridge subsample of the dataset reported by Chen²¹ except for those points annotated by superscripts (1, Reykjanes Ridge)²², (2, Mid-Atlantic Ridge 31° S)²³, and (3, Mid-Atlantic Ridge 39° N)²⁴ (point 3, also R.S. Detrick, personal communication). Note that this last datum is from gravity data which can only determine that the crust here is 2 ± 1 km thicker than neighboring segments to the south which are assumed to be 6 km 'normal' mean crustal thickness). Also shown are a number of theoretical curves, assuming various values of Nu and Z_{cutoff} (all at $T_{cutoff} = 600^\circ\text{C}$). The theoretical curve assuming $Nu = 8$, $T_{cutoff} = 600^\circ\text{C}$, and $Z_{cutoff} = 6$ km best fits the observed transition from axial high to median valley relief as a function of crustal thickness and spreading rate.

depth-integral of the axial yield-strength envelope³. It is a measure of the magnitude of the horizontal tensile stress that can be supported during ridge axis extension. In the lithosphere-stretching hypothesis of the origin of axial topography^{19, 20}, the net horizontal axial force is proportional to the amplitude of the resulting median valley relief. One striking feature of the results shown in Fig. 2 is the existence of a 'threshold' crustal thickness (at a given spreading rate) about which small changes in crustal thickness can produce a dramatic change in axial thermal structure. For example, in the 10 mm yr^{-1} experiments, a change in crustal thickness from 9 to 9.5 km (at $Nu = 8$) leads to a shallowing of the $1,200^\circ\text{C}$ isotherm from 16 to 8 km. This effect can be understood from the nature of the latent heat release in a magma lens—it releases the latent heat of the magma-cumulate phase change for essentially all of the 'cumulate gabbro' crust that resides below the lens. Thus an upward displacement of the level of the magma lens results in more latent heat release at this shallower depth (positive feedback), an effect which would favour a lens at an even shallower depth. This positive-feedback makes these model thermal structures more sensitive to spreading rate or crustal thickness variations than previous models, which treat crustal accretion as an invariant crustal-thickness dike^{7, 19, 20}. Our results are not very sensitive to small amounts (<30%) of latent heat release in the crust beneath the lens. (For example, 25% off-axis latent heat release depresses lens depths by ~ 150 – 300 m as a function of spreading rate; this is an upper

estimate of the latent heat release as >5% intracumulus melt below the lens leads to a larger seismic velocity drop than is observed¹⁵.)

Figure 3 shows the 'threshold' crustal thickness for a steady-state crustal magma lens as a function of spreading rate. (This thickness is defined as the point where the $1,200^\circ\text{C}$ isotherm reaches the base of the crust.) Variations in axial thermal structure are most sensitive at intermediate spreading half-rates of 20 – 30 mm yr^{-1} to small fluctuations in magma supply about the normal rate which leads to a 6-km-thick crust. On this plot we also show observed axial morphology and crustal thickness in areas where there are seismic measurements of near-ridge crustal thicknesses²¹. These measurements support the predictions of our model²². In particular, our model predicts the axial high morphology that is observed for the ~ 9 -km crustal thickness at the Reykjanes Ridge. It also quantitatively agrees with the variations in axial morphology as a function of crustal thickness at the Mid-Atlantic Ridge, 32°S ²³, and 39°N ²⁴.

Gravity and topography data show that slow-spreading ridges have a clear along-axis variation in crustal thickness (that is, the magma supply varies along-axis). In contrast, the much smaller along-axis gravity and topography variation at a fast-spreading ridge can be explained by either (1) a more 2D pattern of upwelling and melting beneath such a ridge, or (2) by a well-connected, temporally persistent magma lens (or low-viscosity zone) that smooths the along-axis crustal structure^{25, 26}. These results suggest that the magmatic heat input associated with a mean crustal thickness of 6 km is sufficient for a quasi-steady-state magma lens to form at fast-spreading ridges. At slower spreading rates (with the same thickness of crust), a continuous lens of this type cannot be a quasi-steady-state feature along an entire segment. Thus, the crustal thickness at slow-spreading ridges is (usually) not strongly smoothed by along-axis flow, and this results in an observed ~ 3 km along-axis crustal thickness variation. The resulting along-axis variations in across-axis thermal structure lead to dramatic variations in axial yield strength (Fig. 2b) and hence in the along-axis relief of median valley topography; variations which this model predicts will be strongly correlated with crustal thickness. This is supported both by observations of strongly-correlated axial Mantle Bouguer Anomaly (a likely measure of crustal thickness) and axial depth profiles at slow- and transitional-spreading ridges^{25, 26}. The origin of magma supply variations is not addressed by our model. Common explanations are lateral variations in mantle temperature or composition beneath the global ridge system²⁷. Both could lead to the same types of variation in magma supply, but may result in discernible variation in crustal composition²⁷. □

Received 11 January; accepted 28 June 1993.

- Macdonald, K. C. in *The Geology of North America: The Western North Atlantic Region* (eds Vogt, P. & Tucholke, B.) 51–68 (Geol. Soc. Am., Boulder, 1986).
- Small, C. & Sandwell, D. T. *J. geophys. Res.* **94**, 17383–17392 (1989).
- Phipps Morgan, J. & Chen, Y. *J. geophys. Res.* **98**, 6283–6298 (1993).
- Purdy, G. M., Kong, L. S. L., Christeson, G. L. & Solomon, S. *Nature* **355**, 815–817 (1992).
- Toomey, D. R., Solomon, S. C. & Purdy, G. M. *J. geophys. Res.* **93**, 9093–9112 (1988).
- Kong, L. S. L., Solomon, S. C. & Purdy, G. M. *J. geophys. Res.* **97**, 1659–1685 (1992).
- Sleep, N. H. *J. geophys. Res.* **80**, 4037–4042 (1975).
- Sinton, J. M. & Detrick, R. S. *J. geophys. Res.* **97**, 197–216 (1992).
- Quick, J. E. & Denlinger, R. P. *J. geophys. Res.* (in press).
- Henstock, T. J., Woods, A. W. & White, R. S. *J. geophys. Res.* **98**, 4143–4162 (1993).
- Sparks, D. W. & Parmentier, E. M. *Earth planet. Sci. Lett.* **105**, 368–377 (1991).
- Harding, A. J. et al. *J. geophys. Res.* **94**, 12163–12196 (1989).
- Kent, G. M., Harding, A. J. & Orcutt, J. A. *Nature* **344**, 650–653 (1990).
- Vera, E. E. et al. *J. geophys. Res.* **95**, 15529–15556 (1990).
- Caress, D. W., Burnett, M. S. & Orcutt, J. A. *J. geophys. Res.* **97**, 9243–9264 (1992).
- Gregory, R. & Taylor, H. P. *J. geophys. Res.* **86**, 2737–2755 (1981).
- Nehlig, P. & Juteau, T. *Tectonophysics* **151**, 199–221 (1988).
- Hooft, E. E. & Detrick, R. S. *Geophys. Res. Lett.* **20**, 423–426 (1993).
- Phipps Morgan, J., Parmentier, E. M. & Lin, J. J. *Geophys. Res.* **92**, 12823–12836 (1987).
- Chen, Y. & Morgan, W. J. *J. geophys. Res.* **95**, 17583–17604 (1990).
- Chen, Y. *J. Geophys. Res. Lett.* **19**, 753–756 (1992).
- Bunch, A. W. H. & Kennett, B. L. N. *Geophys. J. R. astr. Soc.* **61**, 41–166 (1980).
- Toistoy, M., Harding, A. J. & Orcutt, J. A. *EOS* **73**, 495 (1992).
- Detrick, R. S. & Needham, H. D. *EOS* **73**, 569 (1992).
- Phipps Morgan, J. *Rev. Geophys. US National Report to IUGG Supplement*, 807–822 (1991).
- Lin, J. & Phipps Morgan, J. *Geophys. Res. Lett.* **19**, 13–16 (1992).
- Klein, E. M. & Langmuir, C. H. *J. geophys. Res.* **92**, 8089–8115 (1987).
- Huang, P. Y. & Solomon, S. C. *J. geophys. Res.* **93**, 13445–13477 (1988).

A Cambrian gilled lobopod from Greenland

Graham Budd

Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK

THE discovery of several 'Burgess Shale'-like faunas in the Cambrian has added significant new data to the debate about the nature of the 'Cambrian explosion'. Lobopod animals have emerged as a much more important component of the Cambrian fauna than of the Recent^{1,2}. A new lobopod-like animal, *Kerygmachela kierkegaardi* gen. et sp. nov., is now reported from the Lower Cambrian Sirius Passet fauna, north Greenland³. Although it shares features with other recently described lobopods, it differs in possessing lateral lobes along the body with dorsal gill-like structures attached to them, a feature that may indicate a relationship with the enigmatic *Opabinia*⁴. The presence of such a gilled lobopod suggests an evolutionary link between lobopods and the clade consisting of biramous-limbed arthropods^{5,6}. The Cambrian lobopods thus represent a paraphyletic clade from which fully arthropodized taxa were derived. The data from *Kerygmachela*, when considered with

recently published molecular data, may suggest a biphyletic origin for the arthropod grade of organization.

The Lower Cambrian 'Sirius Passet' fauna is mostly known from one locality in Peary Land, north Greenland³. This *lagerstätte* has yielded a rich assemblage of poorly sclerotized fossils, consisting principally of arthropods and sponges. Here I describe a new lobopod-like metazoan, *Kerygmachela kierkegaardi* gen. et sp. nov. from the Sirius Passet fauna that also shares characters with biramous arthropods and two of the most famous Burgess Shale problematica, *Opabinia* and *Anomalocaris*.

Superphylum Lobopodia Snodgrass 1938

Kerygmachela kierkegaardi gen. et sp. nov.

Etymology. *Kerygma*, proclamation (Greek) and *Chela*, claw (Greek), in reference to its flamboyant frontal appendages; *kierkegaardi*, in honour of Søren Kierkegaard, Danish philosopher and resident of Copenhagen, where the type material is housed.

Holotype. MGUH 22.083 from GGU collection 340103 (Geologiske Museum, Øster Volgade, København), part and counterpart of the only complete specimen known. (Figs 1, 2 and 3).

Paratypes. MGUH 22.084, 085, 086 from GGU collection 340103 (Geologiske Museum, Øster Volgade, København). (Figs 1 and 2).

Horizon and locality. Buen Formation (Lower Cambrian), J. P. Koch Fjord, Peary Land, North Greenland.

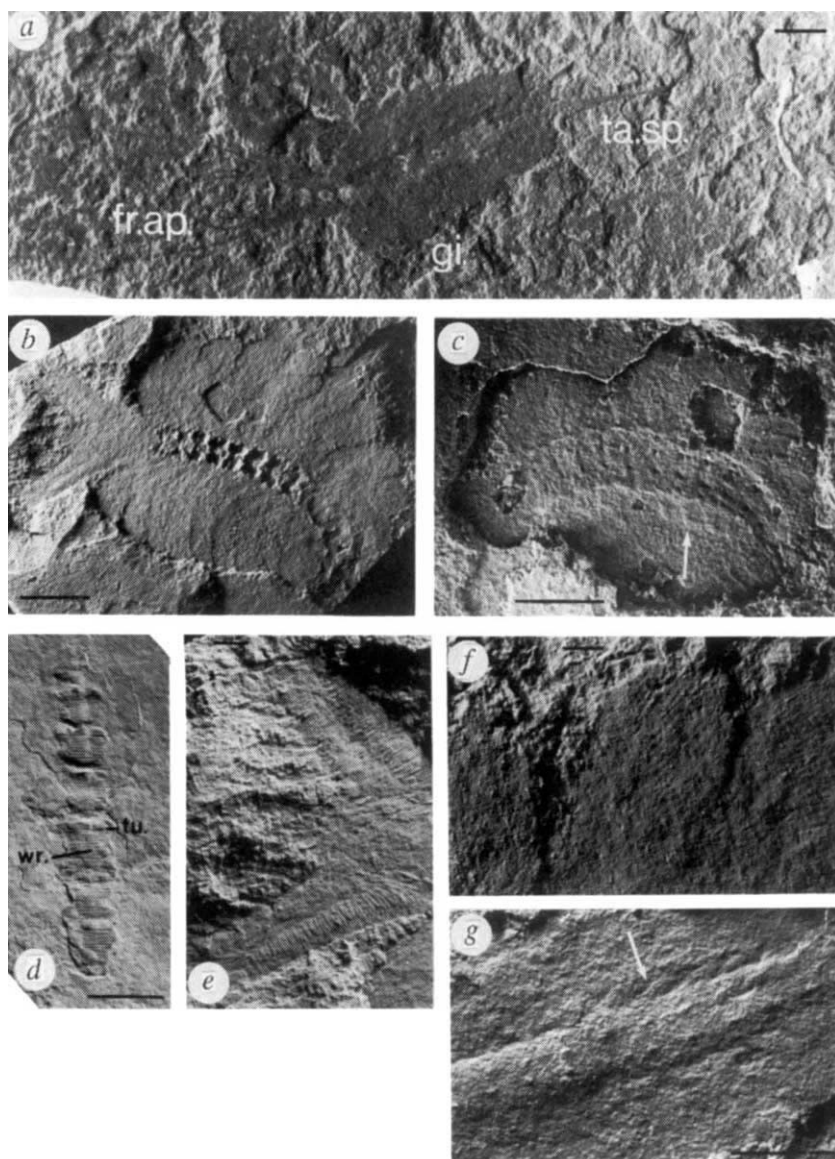


FIG. 1 *Kerygmachela kierkegaardi*. a and g, Complete specimen MGUH 22.083 (holotype). a, Dorsal view; scale bar, 1 cm. g, Close-up of axial region showing details of tubercles and limbs (arrowed); compare *Camera lucida* drawing (Fig. 2c). Scale bar, 1 cm. b, e and f, MGUH 22.084 (paratype). b, Whole specimen; compare *Camera lucida* drawing (Fig. 2a). Scale bar, 1 cm. e, Close-up of cephalic region showing wrinkled and spinose frontal appendages. Scale bar, 2 mm. f, Close-up of lateral lobe showing imbrication of the lobes and the wrinkled, gill-like structures. Scale bar, 1 mm. c, MGUH 22.085 (paratype). Whole specimen; compare *Camera lucida* drawing (Fig. 2b). Limbs arrowed. d, MGUH 22.086 (paratype) showing characteristic axial transverse wrinkling and impressions of tubercles. Scale bar, 1 cm. ta.sp., Tail spines; gi., lateral lobes with gill-like structures; fr.ap., frontal appendage; wr., transverse wrinkling; tu., tubercle.

Diagnosis. Lobopod with stout frontal appendages each bearing four long processes. Thoracic region with eleven segments; axial region includes rows of four tubercles alternating with transversely wrinkled units. Each thoracic segment bears paired lateral gill-like lobes and stout walking appendages. Posterior region is small and trapezoid, bearing two long, segmented tail-spines.

Description. *Kerygmachela kierkegaardi* is known from the part and counterpart of a complete specimen, and about 50 other incomplete specimens. The complete specimen is 17.5 cm long (Fig. 1a). It possesses a well-defined cephalic region with a pair of transversely wrinkled, stout, frontal appendages, each bearing several long slender processes. At the posterior are two long segmented spines. The main body is sub-rectangular in outline; it is composed of 11 segments, each comprising an axial region and a pair of lateral, imbricated lobes that bear wrinkled structures interpreted as gills on the dorsal surface (Figs 1a, b, f and 2a). The axial region of the body shows alternating transverse units of sub-parallel wrinkling and tuberculate rows, each with four tubercles (Figs 1c, d, g and 2b, c). Each segment contains one row of tubercles, flanked on either side by a wrinkled unit. In each tuberculate row, the central two tubercles are larger than the outer ones. In some specimens the wrinkles are deflected as they pass over the dorsal midline of the body, so forming a

distinct crease down the centre of the animal. Early mineralization has preserved some internal structure, including structures interpreted as circularly arranged musculature (Fig. 2a).

Two specimens exhibit sub-triangular bulges projecting laterally from the axial region (Fig. 1c, g; Fig. 2b, c). These bulges are interpreted as stout, lobopod-like limbs. They are preserved in a similar style to other undoubted limbs in Sirius Passet taxa. It is unlikely that the bulges represent the mineralized bases of the gill-like lobes, as the two sets of structures lie at different angles with respect to one another in the two specimens illustrated. In the complete specimen, the bulges connect to the tubercled units of the axial region, a characteristic of other Cambrian lobopods. The overall outline of the axial region and bulges is strikingly similar to that of other lobopods; in particular, their foreshortened and backwards-curved appearance is close to that of the limbs of dorso-ventrally preserved *Aysheaia*⁷.

Although *Kerygmachela* shows a unique combination of morphological features, several of its characters are also possessed in varying degrees by lobopod-like Cambrian animals such as *Luolishania*⁸, *Aysheaia*⁷, *Xenusion*^{2,9}, *Microdictyon*¹⁰, *Onychodictyon*¹¹, *Hallucigenia*¹ as well as the Recent *Peripatus*. These characters include the possession of (1) unsegmented, lobopod limbs; (2) circular body musculature; (3) appendages with spines; and (4) alternating axial units of transverse wrinkling and rows of tubercles¹² (Fig. 3a, b). This last feature is extremely similar to the axial region of *Luolishania*⁸. Transverse wrinkling has been cited as an important synapomorphy in lobopods; it is said to reflect a distinctive arrangement of trans-

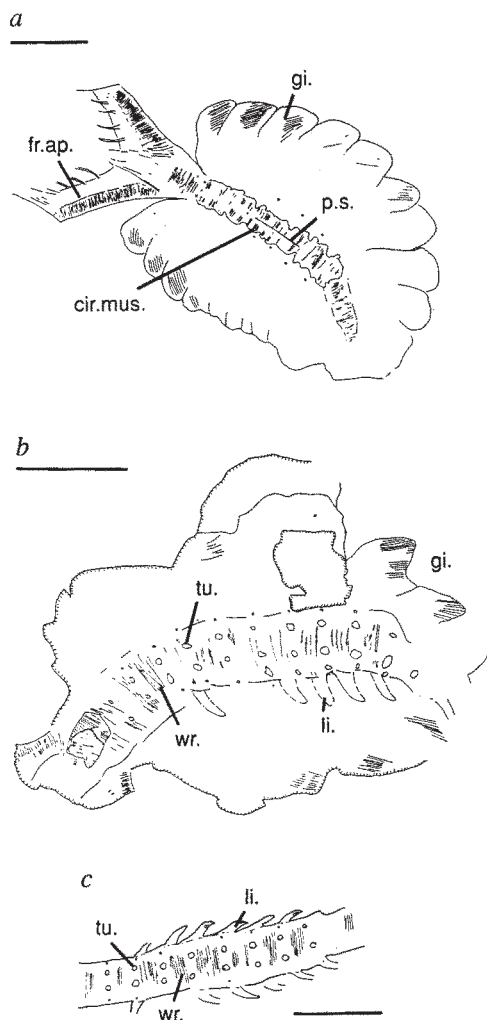


FIG. 2 Camera lucida drawings. a, MGUH 22.084 (paratype) (compare Fig. 1b, e, f). Scale bar, 1 cm. b, MGUH 22.085 (paratype) (compare Fig. 1c). c, MGUH 22.083 (holotype) (compare Fig. 1g). Scale bar, 1 cm. fr.ap., Frontal appendage; gi., lateral lobes with gill-like structures; p.s., inferred pericardial sinus; cir.mus., circular musculature; tu., tubercles; wr., transverse wrinkles; li., limb.

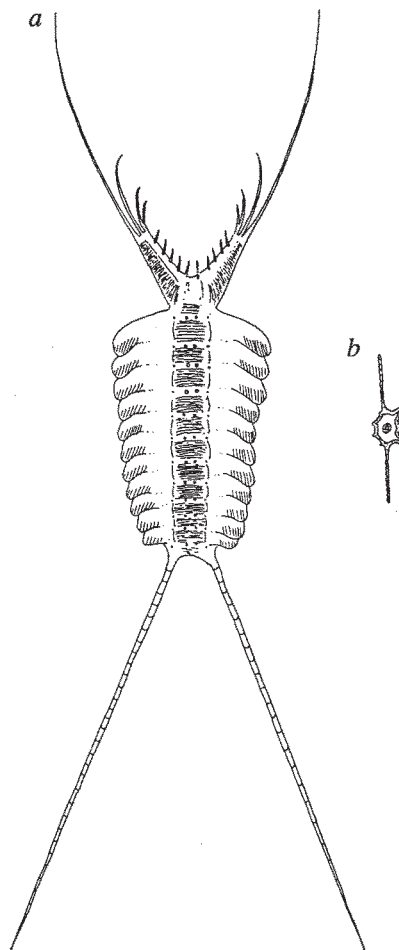
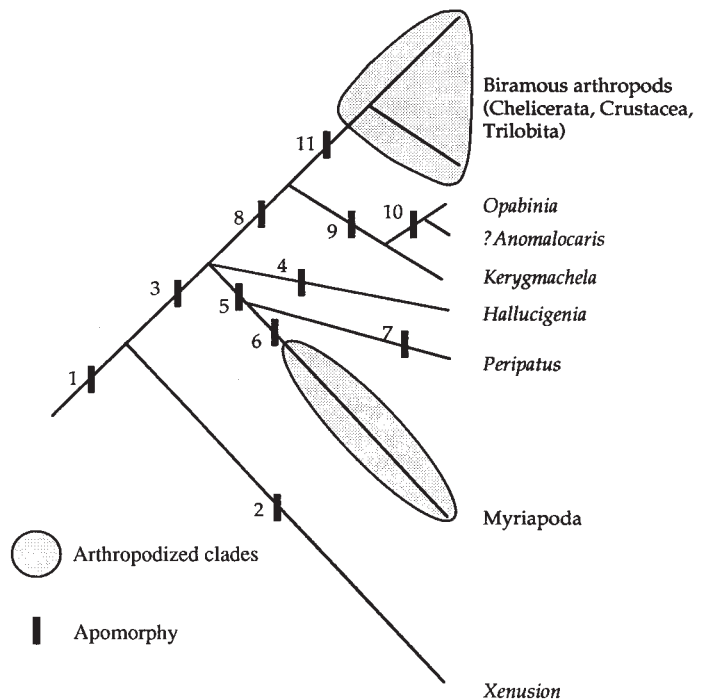


FIG. 3 Reconstructions of *Kerygmachela kierkegaardi*; a, dorsal; b, inferred transverse section.

FIG. 4 Implied phylogeny of 'arthropods' and 'lobopods' based on evidence from *Kerygmachela* and recent molecular and morphological analyses^{9,12,18,19,22-26,28}. The Hexapoda have been excluded from the phylogeny because of uncertainty about their position^{20-22,25,28}, which is unresolved by *Kerygmachela*. The entire clade represents the Lobopodia which includes the sclerotized groups. Expansion of the Onychophora to include all the Cambrian forms converts a monophyletic usage into a paraphyletic one. A primitive position for the myriapods relative to the other arthropodized taxa is suggested by molecular data²³⁻²⁶. 12S RNA data²⁶ places the Onychophora as sister group to the biramous-limbed arthropodized taxa, with the myriapods as an outgroup. A lobopodous origin for the biramous limb, as suggested by *Kerygmachela*, in combination with the 12S RNA data implies an independent arthropodization of the myriapods and the biramous arthropods. This is supported further by the various primitive aspects to the Onychophora such as the presence of cilia and paired segmental nephridia²⁰; these imply that the Onychophora have retained sympleiomorphies from the lobopodous stem-group rather than having become secondarily simplified from an arthropodized myriapod-like ancestor²⁶. Apomorphies: (1) Ostiate dorsal vessel, non-septate haemocoel, chitino-proteinaceous cuticle, deutocerebrum, yolky egg with centrolethical cleavage, uniramous lobopod limbs, dorsal armoured tubercles, undifferentiated head region, ?dorsal wrinkling, ?development of nerve cords from ventral organs, ?sub-cuticular connective tissue endoskeleton; (2) spinose lobopods; (3) differentiation of head from body; (4) elongated spinose armouring; (5) ?whole limb jaws, uniramous antennae, simple eyes; (6) stiffened cuticle with uniramous limbs; (7) oblique body wall musculature, slime papillae; (8) 'biramous' lobopodium, with gills borne on lateral lobes; (9) wrinkled and spinose frontal appendage; (10) loss of dorsal wrinkles and tubercles; (11) two pre-oral appendages, stiffened cuticle with biramous limbs.



verse vascular channels with circular musculature deflected around them¹³. In modern lobopods the wrinkles may be deflected as they pass over the dorsal mid-line, where the musculature is interrupted by the pericardial sinus. The integument therefore shows a tendency to fold along the dorsal mid-line; the presence of such a fold in *Kerygmachela* suggests a similar anatomy (Fig. 2a). The possession of lateral, gill-bearing lobes is, however, a major morphological novelty for this grade of organism. This feature may indicate a relationship with the Burgess Shale problematicum *Opabinia*⁴ (and thus perhaps *Anomalocaris*¹⁴⁻¹⁸) which also possesses dorsal gill-like structures. Other similarities include overall morphology and the wrinkled, spine-bearing frontal appendages. Further, *Opabinia* bears reflective axial sub-triangular axial structures very similar to those interpreted as limbs in *Kerygmachela*. Although these structures have previously been interpreted as gut diverticula⁷, the possibility that they are lobopod limbs needs investigation.

The high-level systematics of arthropods is at present in a state of flux. Little consensus exists on either morphological characters or molecular data¹⁹⁻²⁶; views range from polyphyly to strict monophyly. The position of the lobopods is also uncertain, but most authors regard them as the sister group of the arthropods. *Kerygmachela*, however, provides morphological evidence that the biramous arthropods may be derived from a lobopodous ancestor, giving rise to a taxon like *Marrella*²⁷ that lacks dorsal pleurae and occupies a primitive position in cladistic analyses of biramous arthropods^{6,28}. In this case, the gill-lobe and the lobopod would be united during arthropodization to form the biramous limb. The homology of the lobopod and biramous arthropod limbs implied by *Kerygmachela* suggests that lobopod animals constitute a paraphyletic stem-group assemblage from which arthropodized taxa arose. The usage of the name Onychophora should therefore not be extended to include all Cambrian lobopods; this would convert a monophyletic usage into a paraphyletic one (Fig. 4).

Published molecular data suggests a primitive position for the myriapods within a monophyletic arthropod clade²³⁻²⁶. Further, 12S RNA data²⁶ imply that the uniramous limb is primitive, a

condition which is incompatible with a primitively biramous limb as suggested by *Kerygmachela*. Fully arthropodized taxa may thus have arisen twice from closely related lobopod animals which had the ability to sclerotize their cuticle²² (Fig. 4). In this case then, direct fossil evidence can make a vital contribution to resolving otherwise near-intractable phylogenetic disputes²⁹. □

Received 15 January; accepted 17 June 1993.

1. Ramsköld, L. & Hou, X. *Nature* **351**, 225-228 (1991).
2. Dzik, J. & Krumbiegel, G. *Lethaia* **22**, 169-181 (1989).
3. Conway Morris, S. et al. *Nature* **326**, 181-183 (1987).
4. Whittington, H. B. *Phil. Trans. R. Soc. B271*, 1-43 (1975).
5. Cisne, J. L. *Science* **186**, 13-18 (1974).
6. Briggs, D. E. G. & Fortey, F. A. *Science* **246**, 241-243 (1989).
7. Whittington, H. B. *Phil. Trans. R. Soc. B284*, 165-197 (1978).
8. Hou, X. & Chen, J. *Acta palaeont. sin.* **27**, 208-213 (1989).
9. Dzik, J. in *The Early Evolution of the Metazoa and the Significance of Problematic Taxa* (eds Simonetta, A. M. & Conway Morris, S.) 47-56 (Cambridge, Univ. Press, Cambridge, 1991).
10. Bengtson, S. et al. in *Problematic Fossil Taxa* (eds Hoffman, A. & Nitecki, M. H.) 97-115 (Oxford Univ. Press, New York, 1986).
11. Hou, X. Ramsköld, L. & Bergström, J. *Zoologica Scripta* **20**, 395-411 (1991).
12. Ramsköld, L. *Lethaia* **26**, 443-460 (1992).
13. Robison, R. A. *J. Palaeont.* **59**, 226-235 (1985).
14. Bergström, J. *Lethaia* **19**, 241-246 (1987).
15. Whittington, H. B. & Briggs, D. E. G. *Phil. Trans. R. Soc. B309*, 569-609 (1985).
16. Briggs, D. E. G. & Whittington, H. B. *Lethaia* **20**, 185-186 (1987).
17. Bergström, J. *Lethaia* **20**, 187-188 (1987).
18. Collins, D. *Pal. Soc. Spec. Publ.* **6**, 66 (1992).
19. Willmer, P. G. *Invertebrate Relationships: Patterns in Animal Evolution* 1-400 (CUP, Cambridge, 1990).
20. Boudreaux, H. B. *Arthropod Phylogeny with Special References to Insects* 1-320 (John Wiley & Sons, New York, 1979).
21. Ax, P. *The Phylogenetic System* 1-340 (John Wiley & Sons, Chichester, 1987).
22. Manton, S. M. *The Arthropoda, Habits, Functional Morphology and Evolution* 1-527 (Clarendon, Oxford, 1977).
23. Turbeville et al. *Molec. Biol. Evol.* **8**, 669-686 (1991).
24. Lake, J. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 763-766 (1990).
25. Field, K. G. et al. *Science* **239**, 748-753 (1988).
26. Ballard, J. W. O. et al. *Science* **258**, 1345-1348 (1992).
27. Whittington, H. B. *Bull. geol. Surv. Can.* **209**, 1-24 (1971).
28. Briggs, D. E. G., Fortey, R. A. & Wills, M. *Science* **258**, 1670-1673 (1992).
29. Conway Morris, S. *Nature* **361**, 219-225 (1993).

ACKNOWLEDGEMENTS. I thank J. S. Peel and S. Conway Morris for advice and discussion; R. Fortey, A. Henwood, J. Crampton and N. Butterfield for criticizing an earlier draft; the Director of the Geological Survey of Greenland for hospitality during my work in Copenhagen; and K. Harvey and D. Simons for photography. This work was supported by the NERC, the Carlsberg Foundation, and the Geological Survey of Greenland (GGU).

Hormonally mediated inheritance of acquired characteristics in Mongolian gerbils

Mertice M. Clark, Peter Karpiuk & Bennett G. Galef Jr

Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1, Canada

THE intrauterine position relative to members of the same or opposite sex that a rodent fetus occupies affects both its morphology and behaviour when adult¹⁻¹⁴. Female fetuses that mature between males are androgenized by testosterone crossing fetal membranes^{15,16}, and their phenotypes as adults differ significantly from those of sisters that received less intrauterine exposure to exogenous testosterone¹⁷⁻²⁰. We report here that adult female Mongolian gerbils that gestated between male fetuses produce litters containing a significantly greater proportion of sons than the litters produced by those that gestated between female fetuses. Consequently, daughters delivered by dams that gestated between male fetuses are more likely to have gestated between male fetuses and be androgenized *in utero* than are daughters of dams that gestated between female fetuses. Female gerbils thus tend to inherit the phenotype (either androgenized or not androgenized) of their respective mothers.

In Mongolian gerbils (*Meriones unguiculatus*), as in other rodents, the intrauterine position that a female occupies relative to the sex of its neighbours has a profound impact on its development. Female gerbil fetuses that gestate between two male fetuses (2M females) are exposed to higher levels of testosterone than those that mature between two female fetuses (2F females)²¹. The level of exposure to androgen *in utero* affects both the course of development and the reproductive life history of a female gerbil. For example, androgenized, 2M female gerbils achieve puberty at a significantly greater age than their 2F sisters²² and, when adult, these late-maturing females have only half the lifetime fecundity of their early-maturing sisters²³.

Twenty-five 2M and 25 2F female gerbils were delivered by caesarian section on the last day of their normal gestation²⁴. Each subject female and its littermates were then foster reared by a dam that had produced a litter vaginally on the same day²⁴.

We mated each of our 50 subject females with a sexually proven male when she was 67 days old. On the day of each vaginal delivery an experienced observer who was unaware of the intrauterine position that a dam had occupied as a fetus, examined the anogenital distance of each pup to determine its gender²⁵. We found a significantly greater proportion of males in litters born to 2M females ($57.1 \pm 3.5\%$) than in litters born to 2F females ($43.7 \pm 4.8\%$; Student's *t*-test, $t = 2.73$, d.f. 48, $P < 0.01$). In addition, 2M females from litters that contained a majority of females ($N = 5$) produced a significantly greater percentage of sons throughout their reproductive lifetimes (53.9%) than 2F females from litters that contained a majority of males ($N = 4$; 44.7% ; Mann-Whitney *U* test, $U = 4$, $P < 0.05$, one-tailed). Thus intrauterine position, rather than the sex ratio of litters at birth, predicted the proportion of males in litters. This is important because although the sex ratio of a litter might be influenced by the genotype of its dam, the positions of individual fetuses within uterine horns are random²⁵ and are unaffected by maternal genotype. The finding that a female's intrauterine position (rather than the sex ratio of her natal litter) predicts the sex bias of her lifetime reproductive effort is, therefore, not consistent with the hypothesis that differences in the genotype of 2M and 2F females are transmitted to their respective daughters and cause concordance in the sex bias of litters produced by mothers and daughters.

The probability that a fetus will develop between two fetuses

TABLE 1 Proportion of daughters born to 2M and 2F dams expected to gestate in 2M and 2F intrauterine positions

	Daughters	
	2M	2F
Mothers	0.090	0.052
	2F	0.070
		0.116

of the same or opposite sex is affected by four variables: (1) the sex ratio; (2) size of the litter of which it is a member; and the distribution either (3) within or (4) between uterine horns of the male and female fetuses in a litter.

In previous studies, we determined the number of male and female fetuses carried in the left and right uterine horns of 253 caesarian-delivered gerbil dams on their last day of pregnancy²⁵, and the sex ratios of 265 vaginally delivered litters born in our colony²⁶. The 253 caesarian-delivered dams gestated equal numbers of pups in each uterine horn, but carried a significantly greater proportion of male pups in their right uterine horns and female pups in their left uterine horns²⁵. We observed no deviation from chance in the distribution of male and female pups within each uterine horn. The mean and distribution of the sex ratios of caesarian and vaginally delivered gerbil litters were not significantly different²⁶.

Using these data and the present observations of the sex ratios of our sample of 50 vaginally delivered litters, we computed the expected production of 2M and 2F daughters by both 2M and 2F dams (Table 1). We found the expected probability that the daughter of a 2M mother would be a 2M female was 1.73 times greater than the expected probability that she would be a 2F female. Conversely, the expected probability that the daughter of a 2F mother would be a 2M female was only 0.60 of the expected probability that she would be a 2F female.

The effect of intrauterine position on the sex ratios of litters produced by female gerbils shows that there is a form of hormonally mediated transmission of acquired characteristics. This produces concordance between Mongolian gerbil dams and their daughters in those phenotypic characteristics that are affected by level of exposure to testosterone early in life. □

Received 17 February; accepted 21 May 1993.

- Brown, M. J., Schultz, G. S. & Hilton, F. K. *Endocrinology* **115**, 2318-2323 (1984).
- Clark, M. M., Malenfant, S. A., Winter, D. A. & Galef, B. G. Jr *J. Physiol. Behav.* **47**, 301-305 (1990).
- Clemens, L. G. in *Reproductive Behavior* (eds Montagna, W. & Sadler, W. A.) 25-53 (Plenum, New York, 1974).
- Clemens, L. G., Gladue, B. A. & Coniglio, L. P. *Horm. Behav.* **10**, 40-53 (1978).
- Gandelman, R., vom Saal, F. S. & Reinisch, J. M. *Nature* **266**, 722-724 (1977).
- Kinsley, C., Konen, C., Miele, J., Ghiraldi, L. & Svare, B. *Physiol. Behav.* **36**, 793-799 (1986).
- Kinsley, C., Miele, J., Konen, C., Ghiraldi, L. & Svare, B. *Horm. Behav.* **20**, 7-12 (1986).
- Kinsley, C. et al. *Horm. Behav.* **20**, 201-211 (1986).
- Meisel, R. L. & Ward, I. L. *Science* **213**, 239-241 (1981).
- Vomachka, A. J. & Lisk, R. D. *Horm. Behav.* **20**, 181-193 (1986).
- vom Saal, F. S. *J. Anim. Sci.* **67**, 1824-1840 (1991).
- Wechman, R. Jr, Brown, M. & Hilton, F. *Biol. Reprod.* **33**, 803-807 (1985).
- Zielinski, W. J., vom Saal, F. S. & Vandenbergh, J. G. *Behav. Ecol. Sociobiol.* **30**, 185-191 (1992).
- Tobet, S. A., Dunlap, J. L. & Gerall, A. A. *Horm. Behav.* **16**, 251-258 (1982).
- vom Saal, F. S. & Dahr, M. *Physiol. Behav.* **52**, 163-171 (1992).
- Even, M. D., Dahr, M. & vom Saal, F. S. *J. Reprod. Fertil.* **96**, 709-716 (1992).
- vom Saal, F. S. & Bronson, F. H. *Biol. Reprod.* **19**, 842-853 (1978).
- vom Saal, F. S. & Bronson, F. H. *Science* **208**, 597-599 (1980).
- vom Saal, F. S., Grant, W. M., McMullen, C. W. & Laves, K. S. *Science* **220**, 1306-1309 (1983).
- Clark, M. M., Robertson, R. K. & Galef, B. G. Jr *Dev. Psychobiol.* **26**, 185-194 (1993).
- Clark, M. M., Crews, D. & Galef, B. G. Jr *Physiol. Behav.* **49**, 239-243 (1991).
- Clark, M. M. & Galef, B. G. Jr *Physiol. Behav.* **42**, 15-18 (1988).
- Clark, M. M., Spencer, C. A. & Galef, B. G. Jr *Anim. Behav.* **34**, 551-560 (1986).
- Clark, M. M., Malenfant, S. A., Winter, D. A. & Galef, B. G. Jr *Physiol. Behav.* **47**, 301-305 (1989).
- Clark, M. M. & Galef, B. G. Jr *Dev. Psychobiol.* **23**, 29-37 (1990).
- Clark, M. M., Galef, B. G. Jr & vom Saal, F. S. *Dev. Psychobiol.* **24**, 81-90 (1991).

ACKNOWLEDGEMENTS. We thank M. Bryden for technical assistance and D. Chiszar for his comments on earlier drafts. This research was funded by grants from the Natural Sciences and Engineering Research Council of Canada to M. M. C. and B. G. G.

Rapid sequence evolution of the mammalian sex-determining gene *SRY*

L. Simon Whitfield, Robin Lovell-Badge* & Peter N. Goodfellow

Department of Genetics, Cambridge University, Downing Street, Cambridge CB2 3EH, UK

*Laboratory of Eukaryotic Molecular Genetics, MRC National Institute of Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

IN mammals, induction of male sex determination requires the Y-chromosome gene *SRY*¹. *SRY* encodes a protein with a central 'high mobility group' domain (HMG box) of about 78 amino acids¹⁻³. HMG boxes are found in a wide variety of proteins that bind to DNA with high affinity but differing degrees of sequence specificity⁴. The human *SRY* protein binds to linear DNA with sequence specificity⁵ and to cruciform DNA structures without sequence specificity⁶. The DNA-binding activity of the *SRY* protein resides in the HMG box and mutations in this region are associated with sex reversal in XY females⁶⁻⁸. No function has been ascribed to the portions of the *SRY* protein outside the HMG box. *SRY* belongs to a family of genes that are related by sequence homology within the DNA-binding domain: the genes most similar to *SRY* (>60%) have been named *SOX* genes (*SRY* box genes). None of the known *SOX* genes is homologous to *SRY* outside the HMG-box region. Although *SRY* is an important developmental regulator, its sequence is poorly conserved between species apart from the HMG-box domain. Here we investigate the coding sequence of *SRY* in primates and find that evolution has been rapid in the regions flanking the conserved domain. The high degree of sequence divergence and the frequency of non-synonymous mutations suggest either that the majority of the coding sequence has no functional significance or that directional selection has occurred.

Alignment of *SRY* sequences indicates that sequence conservation is largely confined to the HMG box. Figure 1 presents a comparison of predicted *SRY* amino-acid sequences from eutherians and a metatherian; Fig. 2 shows the sequences of predicted *SRY* proteins from 8 primates. The C-terminal non-box region is heterogeneous in length even within the primates. The marmoset, baboon and orang-utan genes each have an insertion/deletion ('indel') which those of the other primates do not. No two indels are the same but they occur in very similar positions and all maintain the translation frame. In addition, the codon (TAG) that ends the open-reading frame (ORF) in six primate species is altered to TAC in orang-utan and to CAG in marmoset, although the resulting ORF extensions are of different lengths. The predicted mouse *SRY* carboxy terminus is greatly extended (314 residues compared with 70 in human) and features a polyglutamine repeat (ref. 9, and A. Hacker and R.L.B., unpublished data). The N-terminal non-box portion of

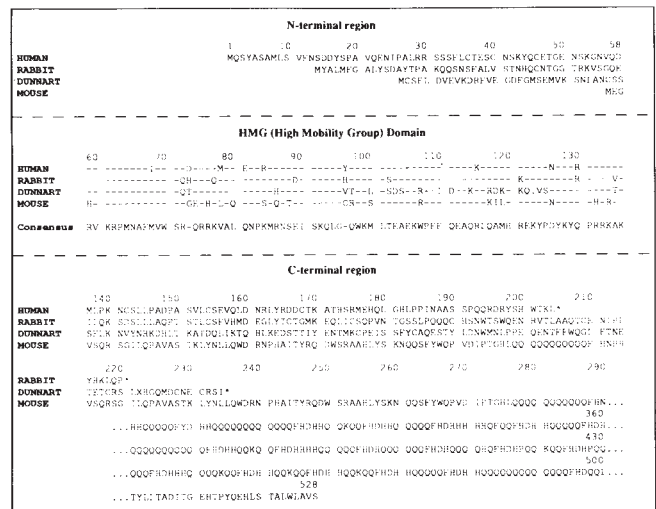


FIG. 1 Comparison of predicted *SRY* protein sequences from Human¹ (primate), Rabbit¹ (lagomorph) and *M. musculus* mouse³ (rodent), all eutherian mammals, and dunnart¹⁹, a marsupial from the dasyurid group. The sequence alignment of the HMG box region was made using PILEUP from the UWGCG package²⁰. Alignments of non-box regions are insufficiently robust to be deemed satisfactory and are not shown. Numbering begins from the initiation methionine of the human sequence and maintains the same numbering sequence through all sequences including mouse.

SRY is homogeneous in length among the primates investigated, although non-primate taxa show heterogeneity for length and it is reduced to three amino acids in the mouse protein.

The rapid divergence of *SRY* sequences outside the HMG box is reflected in the much lower amino-acid identities and similarities seen with pairs of *SRY* protein sequences compared with pairs of MYC and γ -interferon (γ -IFN) sequences (Table 1). Comparison with a broad range of sequences suggests that MYC is conserved whereas γ -IFN is rapidly evolving¹⁰. The absence of conservation of *SRY* non-box regions (95% identity within box, <60% outside it) is in marked contrast to *SRY*'s *SOX* relatives: human and mouse *SOX2* and *SOX3* proteins are both 94% identical over combined box and non-box regions (M. Stevanovic, P.N.G., J. Collignon and R.L.-B., unpublished data).

Comparison of the frequency of DNA point mutations at non-synonymous and synonymous codon positions gives the relative frequency of amino-acid-altering mutations and corrects for the evolutionary separation of the compared species. Figure 2 indicates the location of synonymous and non-synonymous mutations in primate *SRY* sequences, Table 2 shows the levels of non-synonymous mutations per non-synonymous site (K_a) and synonymous mutations per synonymous site (K_s) in the primate

TABLE 1 Percentage amino-acid similarities and identities for human and mouse protein sequences

	Total		N terminus		HMG box		C terminus	
	Identity	Similarity	Identity	Similarity	Identity	Similarity	Identity	Similarity
Human/marmoset <i>SRY</i>	74	83	62	76	95	98	59	74
Human/marmoset cMYC	96	97						
Human/marmoset IFN- γ	89	96						
Human/mouse IFN- γ	41	62						
Human/mouse <i>SOX3</i>	94	96						

Calculations were done using BESTFIT from the UWGCG suite of programs¹⁹. For human *SRY* protein, N terminus refers to amino acids 1-58, HMG box is from 59-135, and C terminus from 135 onwards. Human, mouse and marmoset non-*SRY* sequences were obtained from the EMBL sequence database, except for *SOX3* sequences (M. Stevanovic, P.N.G., J. Collignon and R.L.-B., unpublished data).

SRY genes. When seven non-human primate *SRY* sequences are compared with human *SRY*, K_a/K_s lies between 0.32 and 1.88 (although the standard error on the estimates for closely related species are high). K_a/K_s for the most diverged pair of species (human to marmoset) is 0.47. This value is markedly higher than that observed in a range of coding sequences¹⁰: the average for 42 sequences was 0.189, the highest values coming from the interferons (for example, γ -IFN was 0.3). Table 2 includes values of K_a , K_s and K_a/K_s for the three regions of the primate *SRY* proteins when compared with human *SRY*. For most pairings, K_a/K_s for N- and C-terminal domains are either close to unity or undefined because the sequence pairs show non-synonymous differences but no synonymous differences. It appears, moreover, that the high values of K_a/K_s are not simply an effect of the peculiar genetic environment of the sex-specific region of the Y chromosome: K_a/K_s for the comparison of human *ZFY* to mouse *Zfy-1* is 0.19. It should be noted that the different expression patterns of *ZFY* and *Zfy-1* indicate that their functions may have diverged^{11,12}, involving a degree of selection for non-synonymous mutations.

The observed values of K_a/K_s reflect the relative strengths of opposing evolutionary forces, functional constraints tending to conserve sequence identity versus divergence caused by neutral drift and selection. A K_a/K_s ratio significantly greater than one suggests that directional selection has occurred. If it is assumed that there is a function for the non-box coding regions of *SRY*, the high values of K_a with respect to K_s imply that the observed divergence between the primate *SRY* sequences may be due to selection. The observation that the ratio between K_a and K_s may vary between different parts of the primate tree also suggests that divergence has not been caused by neutral drift. The alternative hypothesis is that these regions have been retained through the primate lineage without any requirement for sequence conservation, implying that the *SRY* protein is, in functional terms, the HMG box alone. The exact functional relationships

TABLE 2 Synonymous substitutions per synonymous site [$(K_s) \times 100 \pm \text{s.e.} \times 100$], non-synonymous substitutions per non-synonymous site [$(K_a) \times 100 \pm \text{s.e.} \times 100$] and their ratio (K_a/K_s) for human to primate *SRY* sequence pairs

		SRY whole	N terminus	HMG box	C terminus
$K_s \times 100$	Chimp	0.8 ± 0.8	3 ± 3.1	0 ± 0	0 ± 0
	Pyg chimp	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	Gorilla	2.5 ± 1.5	3.1 ± 3.1	2.3 ± 2.3	2.3 ± 2.3
	Orang-utan	2.5 ± 1.5	3.1 ± 3.1	4.6 ± 3.4	0 ± 0
	Gibbon	5.9 ± 2.3	3 ± 3.1	6.9 ± 4.1	7.5 ± 4.5
	Baboon	14.3 ± 3.7	21.7 ± 10.3	9.6 ± 4.9	15.1 ± 6.4
	Marmoset	27.6 ± 5.5	43.1 ± 14.9	18.7 ± 7	26.7 ± 9.2
$K_a \times 100$	Chimp	1.5 ± 0.6	3.8 ± 1.7	0.5 ± 0.5	0.6 ± 0.6
	Pyg chimp	2.3 ± 0.7	4.6 ± 1.9	1.1 ± 0.8	1.9 ± 1.1
	Gorilla	0.8 ± 0.4	3 ± 1.5	0 ± 0	0 ± 0
	Orang-utan	2.5 ± 0.7	3 ± 1.5	0.5 ± 0.5	4.5 ± 1.7
	Gibbon	3.6 ± 0.9	5.3 ± 2	1.1 ± 0.8	5.2 ± 1.9
	Baboon	5.4 ± 1.1	7.9 ± 2.6	1.1 ± 0.8	8.6 ± 2.4
	Marmoset	12.9 ± 1.8	19 ± 4.2	2.3 ± 1.1	21.2 ± 4.1
K_a/K_s	Chimp	1.88	1.27	n/0	n/0
	Pyg chimp	n/0	n/0	n/0	n/0
	Gorilla	0.32	0.97	0	0
	Orang-utan	1	0.97	0.11	n/0
	Gibbon	0.61	1.77	0.16	0.69
	Baboon	0.38	0.36	0.11	0.57
	Marmoset	0.47	0.44	0.12	0.79

Calculations were executed according to ref. 10 using the program LWL produced by K. Wolfe. The HMG box and non-box regions are as defined for Table 1.

of the box and non-box regions are being studied in transgenic animals.

Rapid divergence by natural selection at the molecular level has only been demonstrated in few cases, typically involving intergenomic conflict¹³⁻¹⁶ such as occurs between host and pathogen: the binding pocket of the major histocompatibility

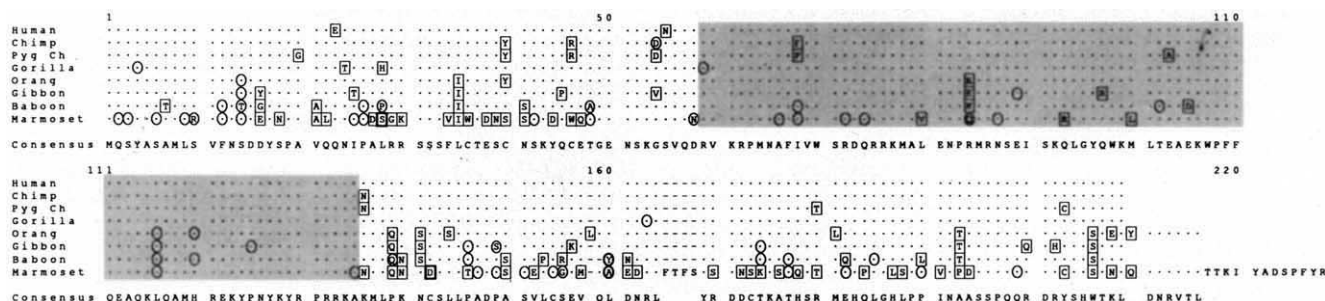


FIG. 2 Alignment of predicted *SRY* protein sequences from primates showing the location of non-synonymous and synonymous DNA point mutations. The HMG box is shaded. Dots indicate agreement with the consensus amino-acid sequence shown and dashes denote gaps. Boxes indicate the location of non-synonymous mutations with the resultant amino-acid change shown. Circled dots indicate codons containing synonymous mutations. Circled letters denote codons containing a synonymous and a non-synonymous mutation. Bold boxes and circles denote 2 of the relevant mutation. In all cases the minimum number of non-synonymous mutations is assumed. Alignment and consensus were produced using PILEUP and PRETTY from UWGCG²⁰.

METHODS. Male and female common chimpanzee (*Pan troglodytes*), pygmy chimpanzee (*Pan paniscus*) and orang-utan (*Pongo pygmaeus*) DNA samples were prepared from lymphoblastoid cell lines Colin, Halpha, Bosondjo, Laura, Chantek and Gucchi (in order) provided by S. Marsh and P. Parham. Male and female gorilla (*Gorilla gorilla*) DNA samples were provided by J. Hearn, male gibbon (*Hylobates lar*) DNA by H. Stanley, female gibbon by D. Bailey, baboon (*Papio spp.*) male and female DNA samples by H. Stanley, and S. Rastan provided male and female common marmoset (*Callithrix jacchus*) tissue samples from which genomic DNA was prepared²¹. Primate *SRY* sequences were amplified from genomic DNA using the polymerase chain reaction (PCR);

33 cycles of 92 °C for 30 s, 50 °C for 30 s, 72 °C for 45 s) with *Taq* polymerase (Promega) and primers SW2 (CTTGAGAATGAATACATTGTCAGGG) and SW3B (biotin-AGGTCTTTGTAGCCAATGTTACCCG). Male and female DNA samples from each species were tested simultaneously and a single, male-specific band was seen for each species. Alkali blotting onto Hybond N+ (Amersham) and hybridization with a human *SRY* probe confirmed that the single bands were indeed Y-specific. Primate *SRY*s were sequenced directly from the PCR product. The male-specific band from each species was gel-purified and ~2 ng product reamplified (25 cycles as described) in 50-μl reactions using 5 pmol SW2 and SW3B. The use of a single biotinylated oligonucleotide permitted the preparation of single-stranded DNA from the PCR product; double-stranded DNA (45 μl PCR product) was bound to 10 μl streptavidin-coated magnetic beads (Dyna) and the non-biotinylated strand removed by alkali denaturation. Non-biotinylated and bead-bound biotinylated strands were sequenced by the dideoxy chain termination method using Sequenase 2.0 (USB), in accordance with manufacturer's instructions. Phylogenetic analyses were performed using DNABOOT and DNAML from PHYLIP²², confirming that the sequences obtained did indeed form a molecular lineage appropriate to the currently accepted phylogeny of the primates.

complex class I antigens^{14,15} and the T-cell epitopes of the circumsporozoite protein of the malaria parasite *P. falciparum*¹⁶, for example, have both undergone directional selection.

It is not clear how the high number of amino-acid substitutions might have been subject to advantageous selection in primate *SRY* genes. Should a form of genetic conflict prove to be involved, it may not be intergenomic, but intragenomic or even intrachromosomal. It is clear, however, that if *SRY* is periodically subject to strong selective forces then it will be a major

driving force in Y-chromosome evolution, causing rapid substitution of Y-chromosome variants in strong linkage disequilibrium with the *SRY* variant undergoing substitution. The occurrence of such sweeps may contribute to the apparent absence of polymorphism on the Y chromosome^{17,18}. Moreover, given that reduced or inappropriate *SRY* activity can cause sex reversal, the rapid evolution of *SRY*, producing populations with different *SRY* sequences, could be a significant cause of reproductive isolation. □

Received 14 April; accepted 21 May 1993.

1. Sinclair, A. H. et al. *Nature* **346**, 240–244 (1990).
2. Su, H. & Lau, Y.-F. C. *Am. J. Hum. Genet.* **52**, 24–38 (1993).
3. Gubbay, J. et al. *Nature* **346**, 245–250 (1990).
4. Ner, S. S. *Curr. Biol.* **2**, 208–210 (1992).
5. Harley, V. R. et al. *Science* **255**, 453–456 (1992).
6. Ferrari, S. et al. *EMBO J.* **11**, 4497–4506 (1992).
7. Berta, P. et al. *Nature* **348**, 448–450 (1990).
8. Jager, R. J., Anvret, M., Hall, K. & Scherer, G. *Nature* **348**, 452–454 (1990).
9. Gubbay, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7953–7957 (1992).
10. Li, W. H., Wu, C. I. & Luo, C. C. *Molec. Biol. Evol.* **2**, 150–174 (1985).
11. Koopman, P., Gubbay, J., Collignon, J. & Lovell-Badge, R. *Nature* **342**, 940–942 (1989).
12. Palmer, M. S., Berta, P., Sinclair, A. H., Pym, B. & Goodfellow, P. N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1681–1685 (1990).
13. Hill, R. E. & Hastie, N. D. *Nature* **326**, 96–99 (1987).
14. Hughes, A. L. & Nei, M. *Nature* **335**, 167–170 (1988).

15. Hughes, A. L., Ota, T. & Nei, M. *Molec. Biol. Evol.* **7**, 515–524 (1990).
16. Hughes, A. L. *Genetics* **127**, 345–353 (1991).
17. Malaspina, P. et al. *Ann. hum. Genet.* **54**, 297–305 (1990).
18. Ellis, N. et al. *Nature* **344**, 663–665 (1990).
19. Foster, J. W. et al. *Nature* **359**, 531–533 (1992).
20. GCG Program Manual for the GCG Package (575 Science Drive, Madison, Wisconsin 53711, USA, 1991).
21. Johns, M. B. & Paulus-Thomas, J. E. *Analyt. Biochem.* **180**, 276–278 (1989).
22. Felsenstein, J. *Cladistics* **5**, 164–166 (1989).

ACKNOWLEDGEMENTS. We thank J. A. Barrett, W. A. Amos, J. R. Hawkins, C.-I. Wu, A. Hughes and our colleagues for various comments, advice and computer programs, and the providers of primate cell lines and DNA samples. L.S.W. holds a Bursarship from the Imperial Cancer Research Fund. P.N.G. is supported by the Wellcome Trust and R.L.-B. is supported by the UK MRC.

Rapid evolution of the sex determining locus in Old World mice and rats

P. K. Tucker & B. L. Lundrigan

Museum of Zoology and Department of Biology,
University of Michigan, Ann Arbor, Michigan 48109, USA

THE Y chromosome-linked sex determining locus (*Sry*) responsible for testis determination in mammals^{1–5} contains a DNA-binding motif (HMG box) that is conserved across species of marsupial and placental mammals (infraclasses Metatheria and Eutheria, respectively)^{1,4,6,7}. But little to no sequence similarity is observed in flanking sequences between these two infraclasses, or among orders within each infraclass. We investigated the rate and pattern of evolution for the coding sequence of *Sry* in Old World mice and rats (subfamily Murinae). We found typical rates of synonymous substitution throughout the gene, but high rates of non-synonymous substitution, especially in the C-terminal (non-HMG box) region, when compared to other genes⁸. This region is also characterized by a frame-shift mutation and variation in a trinucleotide repeat motif. These data suggest that the non-box region is either functionally unconstrained or has undergone species-specific adaptive divergence.

The mouse *Sry* transcript, like human *SRY*⁹, probably consists of a single exon with an open reading frame consisting of a central HMG (high mobility group) DNA-binding domain flanked by N-terminal and C-terminal regions (A. Hacker and R. Lovell-Badge, manuscript in preparation). We have compared a portion of the *Sry* transcript (86 base pairs (bp) of the 5'-untranslated region, the N-terminal region, the HMG box and the C-terminal region) in seven species of Old World mice and rats belonging to five genera within the subfamily Murinae, family Muridae (Fig. 1). Sequence alignment was possible for all regions except a portion of the C-terminal region (amino-acid positions 144–366) which, because of a highly variable trinucleotide repeat, could not be unambiguously aligned. The C-terminal region is also characterized by a frame-shift mutation at amino-acid position 138. A phylogenetic analysis of our aligned sequence (B.L.L. and P.K.T., manuscript in preparation) indicates that the frame-shift mutation results from the insertion

TABLE 1 Per cent sequence divergence between *Mus musculus* and six species of Old World mice and rats for aligned regions of the *Sry* locus calculated by the Jukes and Cantor method²²

Species	Total 515 bp	5' untranslated 86 bp	N-terminus/ HMG box 243 bp	C-terminus 186 bp
Ms	1.7	1.2	0.4	3.8
Mp	5.8	3.5	2.5	11.3
Mh	7.0	8.1	3.3	11.3
Ha	8.2	4.7	4.5	14.5
Sl	10.3	12.8	4.9	16.1
Re	11.7	12.8	6.2	18.3

Aligned regions include a portion of the 5' untranslated region, the N-terminal region and HMG box (amino-acid positions 1–81), and a portion of the C-terminal region (82–143). Species are as described in Fig. 1.

of a thymine in the ancestral lineage giving rise to *Mastomys hildebrandtii* and *Hylomyscus alleni*.

The trinucleotide repeat varies in length among species, from 60 bp in *Hylomyscus alleni* to 666 bp in *Mus musculus*. The predominant repeat motif is CAG in each species. But in each species this motif is interrupted at irregular intervals by single base pair changes and, in *Mus musculus*, by a TTC CAT motif. The frame-shift mutation, coupled with variation in both base pair composition and length of the trinucleotide repeat, produce extensive interspecific variation in amino-acid composition in this region of the gene.

Sequence divergence (in per cent) between *Mus musculus* and each of the other species is calculated for the various regions of the aligned sequence (Table 1). In every comparison, the N terminus/HMG box is the most conserved region of the gene and the C terminus is the least conserved. The pattern of synonymous and non-synonymous substitutions for these same comparisons is shown in Table 2. The ratio of non-synonymous to synonymous substitutions is greater in the C terminus than in the N terminus/HMG for all comparisons except that between *Mus musculus* and *Mastomys hildebrandtii*. This exception is probably due to the small number of total substitutions in the N-terminus/HMG region between these two species.

Our comparison between *Mus musculus* and *Rattus exulans* *Sry* sequence was examined in light of comparative data for the same genera from 28 other gene sequences⁸. The rates of

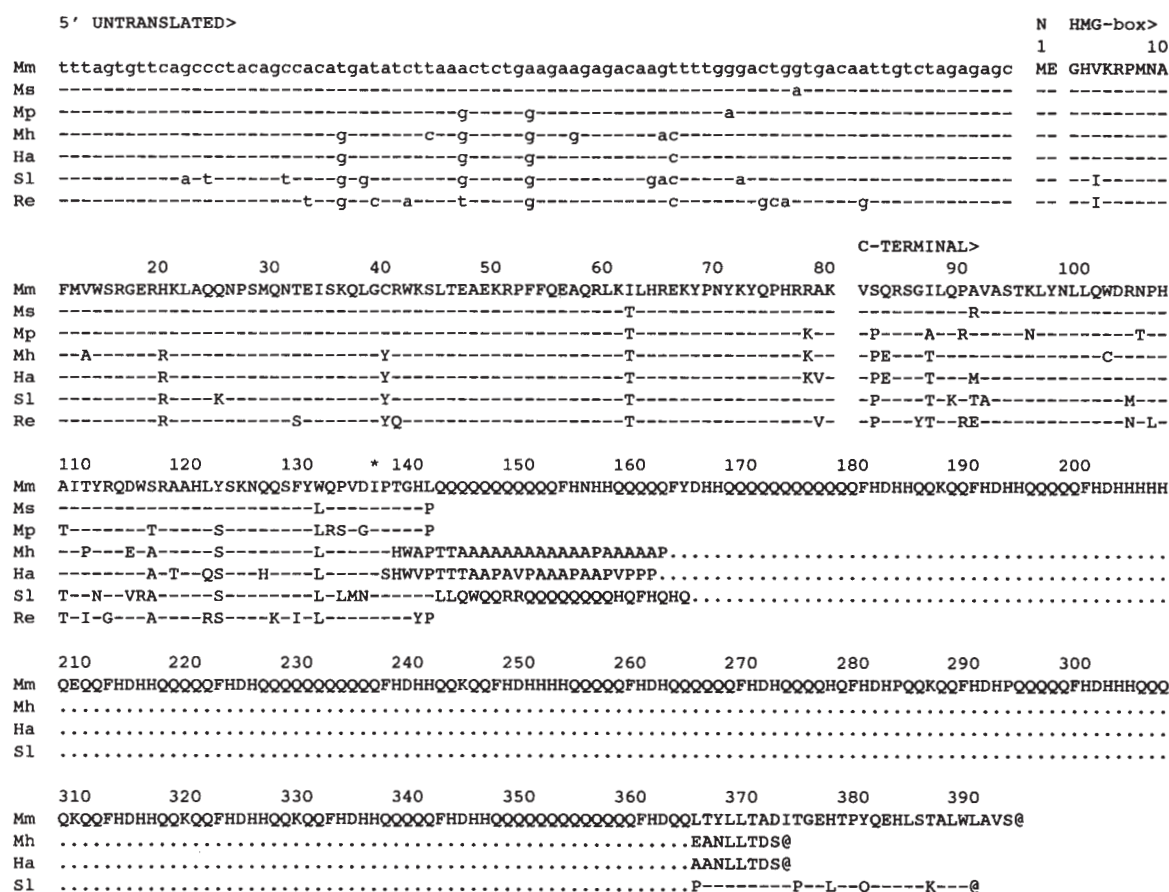


FIG. 1 Nucleotide and amino-acid sequences (single-letter code) for Sry from seven species of Old World mice and rats (subfamily Murinae, family Muridae). Aligned sequences include 86 bp of the 5' untranslated region, the N-terminal region (amino-acid positions 1–2), the HMG box (3–81) and a portion of the C-terminal region (82–143). A highly variable trinucleotide repeat motif in the C-terminal region (144–366) could not be unambiguously aligned at either the nucleotide or amino-acid level. The remainder of the C-terminal region (367–395) could not be aligned at the amino-acid level between Mm and Mh/Ha due to a single base pair insertion (*) at position 138 causing a frame-shift mutation. Dashes indicate sequence identity in aligned regions. Dots are gaps in the sequence. @, Stop codon; Mm, *Mus musculus*; Ms, *Mus spretus* (Spain); Mp, *Mus pahari* (Thailand); Mh, *Mastomys hildebrandtii* (Kenya); Ha, *Hylomyscus alleni* (Gabon); Sl, *Stochomys longicaudatus* (Gabon); Re, *Rattus exulans* (Philippines). Data for amino-acid positions 144–395

are missing for three species (Ms, Mp and Re). PCR amplifications of this region in Ms and Mp indicate that it is roughly equal in length to that for Mm. The repeat region in Re is considerably shorter. Combinations of forward (5'AGATCTTGATTTTATGTTTC3', 5'GGCCATGTCAGCGCCCATG3', 5'TCCTACACAGAGAGAAATACC3') and reverse (5'TGCAGCTCTACTCCAGTCTTG3', 5'CACAGTGTGCTGCTGTAGTA3', 5'GCTGTTGCTGCTTTGTGCTAGC3') primers were used in amplification experiments²⁴. Sequences were aligned to the sequence from mouse strain 129²⁵, which has an Asian *Mus musculus* Y chromosome²⁶, using the SS2 algorithm²⁷, available in the Eugene package of sequence analysis programs. Alignments were subsequently modified by visual inspection. 5' untranslated, N-terminus/HMG box, and C-terminal regions correspond to nucleotide positions 8,217–8,303, 8,304–8,546 and 8,547–9,488, respectively, from the 14.5-kb genomic fragment (L741) containing the Sry gene isolated from mouse strain 129²⁵.

TABLE 2 Synonymous (ds) ± s.e. and non-synonymous (dn) ± s.e. substitutions per 100 sites and their ratio (dn/ds) for aligned regions of the Sry locus

Species	Total			N terminus/HMG			C terminus		
	ds	dn	dn/ds	ds	dn	dn/ds	ds	dn	dn/ds
Ms	3.3 (1.87)	1.5 (0.66)	0.45	0.0	0.5 (0.52)	—	6.9 (3.86)	2.8 (1.38)	0.41
Mp	11.8 (3.33)	4.8 (1.17)	0.41	8.3 (3.98)	1.0 (0.73)	0.12	15.4 (5.37)	10.0 (2.52)	0.64
Mh	16.7 (3.88)	6.5 (1.35)	0.39	6.2 (3.44)	2.6 (1.15)	0.41	28.5 (6.82)	11.6 (2.69)	0.40
Ha	22.7 (4.35)	7.5 (1.44)	0.33	12.4 (4.74)	2.7 (1.17)	0.21	33.3 (7.20)	13.8 (2.88)	0.41
Sl	20.9 (4.21)	6.8 (1.38)	0.33	14.3 (4.99)	2.6 (1.16)	0.18	28.4 (6.79)	12.3 (2.76)	0.43
Re	22.1 (4.31)	8.6 (1.53)	0.39	15.4 (5.18)	3.9 (1.40)	0.25	29.3 (6.86)	14.8 (2.98)	0.50

Comparisons are between *Mus musculus* and six species of Old World mice and rats. The number of synonymous and non-synonymous sites compared for each analysis are as follows. Total, 91.0 to 93.5 synonymous and 332.50 to 334.75 non-synonymous; N terminus/HMG, 48.0 to 49.0 synonymous and 188.58 to 192.0 non-synonymous; C terminus, 42.58 to 45.33 synonymous and 140.67 to 143.42 non-synonymous. Calculations followed the method of Nei and Gojobori²³ using the program NAG provided by A. Hughes. Species are as described in Fig. 1; N-terminus/HMG and C-terminal regions of the gene are as described in Table 1. Comparisons between Mm and Ms, Mm and Mp, and Mm and Mh for the N terminus/HMG are based on a small number of substitutions and may not be statistically meaningful.

synonymous substitution for the N-terminus/HMG and C-terminal regions fall well within the range of synonymous substitutions for these 28 genes (median = 21.6, $r = 9.4-60.2$). In contrast, the rates of non-synonymous substitution for the N-terminus/HMG region and the combined N-terminus/HMG and C-terminal regions are among the highest of the 28 genes (median = 1.3, $r = 0-11.6$). Indeed, the C-terminal region alone shows a higher rate of non-synonymous substitution than any of these 28 genes.

The pattern of substitution for *Sry* relative to other genes was assessed by examining the ratio of non-synonymous to synonymous substitution. The dn/ds for the N-terminus/HMG box, C-terminal and combined regions are among the highest of all gene comparisons (median = 0.07, $r = 0.01-0.58$). Taken together these data suggest that, relative to other genes, *Sry*, and in particular the C-terminal region, is undergoing rapid divergence in amino-acid sequence.

This rapid sequence divergence could be explained in one of two ways: either the *Sry* gene (in particular the C-terminal region) is not functionally constrained, or it is undergoing species-specific adaptive divergence by a process of positive darwinian selection. Circumstantial evidence against the hypothesis of relaxed selection in the C-terminal region includes the following observations. First, although both the human and mouse *SRY* proteins bind the same consensus DNA sequence *in vitro*¹⁰, XX mice transgenic for a 25 kilobase (kb) genomic fragment containing the human *SRY* gene, express the gene, but are not sex-reversed^{5,10}. One possible explanation for this result is that the highly variable region outside the DNA-binding domain is important for the functioning of *Sry*. Second, the trinucleotide repeat in the C-terminal region of *Sry* may affect expression of the gene. Several human genetic disorders result from the presence of trinucleotide repeats either within the coding portion of a gene^{11,12} or in the 5' or 3' untranslated regions¹³⁻²¹. In every case, variation in the length of the repeat is thought to affect the normal expression of the gene or the structure of the protein product. Likewise, variation in the repeat may affect structure and function of *Sry* in a species-specific way. Ultimately, the evolutionary significance of rapid amino-acid sequence divergence in *Sry* may be best understood using a transgenic approach to assess how different naturally occurring variants of the gene, or portions thereof, from one species, function in a second closely related species. □

Isolation of a Miller-Dieker lissencephaly gene containing G protein β -subunit-like repeats

Orly Reiner*, Romeo Carrozzo*, Ying Shen*, Manfred Wehnert*, Fabrizia Faustinella*, William B. Dobyns†, C. Thomas Caskey*‡§ & David H. Ledbetter*

* Institute for Molecular Genetics, and † Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030, USA

‡ Division of Pediatric Neurology, Departments of Neurology and Pediatrics, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA

LISSENCEPHALY (agyria-pachygyria) is a human brain malformation manifested by a smooth cerebral surface and abnormal neuronal migration^{1,2}. Identification of the gene(s) involved in this disorder would facilitate molecular dissection of normal events in brain development³. Type 1 lissencephaly occurs either as an isolated abnormality or in association with dysmorphic facial appearance in patients with Miller-Dieker syndrome^{4,5}. About 15% of patients with isolated lissencephaly and more than 90% of patients with Miller-Dieker syndrome have microdeletions in a critical 350-kilobase region in chromosome 17p13.3 (ref. 6). These deletions are hemizygous, so haplo-insufficiency for a gene in this interval is implicated. Here we report the cloning of a gene (*LIS-1*, lissencephaly-1) in 17p13.3 that is deleted in Miller-Dieker patients. Non-overlapping deletions involving either the 5' or 3' end of the gene were found in two patients, identifying *LIS-1* as the disease gene. The deduced amino-acid sequence shows significant homology to β -subunits of heterotrimeric G proteins, suggesting that it could possibly be involved in a signal transduction pathway crucial for cerebral development.

During a study to identify genes containing β -transducin-like repeats⁷, degenerate polymerase chain reaction (PCR) primers were designed for conserved amino-acid sequences⁸ and used to amplify DNA from a human fetal brain complementary DNA library. The product of this amplification, when used to screen the same library, identified three cDNAs. One corresponding to the known protein 12.3 (ref. 9) and the other two were new cDNAs. The sequence of these two cDNAs was very similar, but PCR analysis of a panel of monochromosomal hybrids mapped one clone to chromosome 2 and the other to chromosome 17 (data not shown). Multiple isolates of cDNAs from chromosomes 2 and 17 were obtained from brain and kidney libraries, but here we study the gene from chromosome 17.

Several cDNA clones mapping to chromosome 17 were isolated and sequenced (Fig. 1). The presence of at least two 5' sequences is indicative of alternative splicing. Regional mapping to a small panel of somatic cell hybrids containing different segments of chromosome 17 (ref. 10) localized the cDNAs to distal 17p (p12-pter). The cDNAs were next analysed in somatic cell hybrids containing chromosomes 17 isolated from patients with Miller-Dieker lissencephaly syndrome (MDS)⁶. Two cDNA clones, representing 3' (clone 6-1) and 5' (clone 8-1) portions of the gene, were hybridized to a panel of somatic cell hybrids containing a single chromosome 17 from several MDS patients. Representative results are shown in Fig. 2a, b. Hybrid KCB4 contains the deleted chromosome from patient MDS-19 (ref. 11), one of four MDS patients whose breakpoints define the centromeric boundary of the critical region¹². The Southern blot results indicate that there is a partial deletion involving only

Received 24 May; accepted 22 June 1993.

1. Sinclair, A. H. et al. *Nature* **346**, 240-244 (1990).
2. Berta, P. et al. *Nature* **348**, 448-450 (1990).
3. Jager, R. J., Anvret, M., Hall, K. & Scherer, G. *Nature* **348**, 452-454 (1990).
4. Gubbay, J. et al. *Nature* **346**, 245-250 (1990).
5. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. *Nature* **351**, 117-121 (1991).
6. Foster, J. W. et al. *Nature* **359**, 531-533 (1992).
7. Whitfield, L. S., Lovell-Badge, R. & Goodfellow, P. N. *Nature* **364**, 713-715 (1993).
8. O'Huigin, C. & Li, W.-H. *J. molec. Evol.* **35**, 377-384 (1992).
9. Su, H. & Lau, Y.-F. C. *Am. J. hum. Genet.* **52**, 24-38 (1993).
10. Lovell-Badge, R. *Phil. Trans. R. Soc. B339*, 159-164 (1993).
11. Laspadula, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fishbeck, H. *Nature* **352**, 77-79 (1991).
12. The Huntington's Disease Collaborative Research Group *Cell* **72**, 971-983 (1993).
13. Fu, Y.-H. et al. *Cell* **67**, 1047-1058 (1991).
14. Kremer, E. J. et al. *Science* **252**, 1711-1714 (1991).
15. Verkerk, A. J. M. H. et al. *Cell* **65**, 905-914 (1991).
16. Aslanidis, C. et al. *Nature* **355**, 548-551 (1992).
17. Brook, J. D. et al. *Cell* **68**, 799-808 (1992).
18. Buxton, J. et al. *Nature* **355**, 547-548 (1992).
19. Fu, Y.-H. et al. *Science* **255**, 1256-1259 (1992).
20. Harley, H. G. et al. *Nature* **355**, 545-546 (1992).
21. Mahadevan, M. et al. *Science* **255**, 1253-1255 (1992).
22. Jukes, T. H. & Cantor, C. R. in *Mammalian Protein Metabolism* Vol. 3 (ed. Monro, H. N.) 21-120 (Academic, New York, 1969).
23. Nei, M. & Gojobori, T. *Molec. Biol. Evol.* **3**, 418-426 (1986).
24. Tucker, P. K. & Lundrigan, B. L. *Meth. Enzym.* **224**, 517-525 (1993).
25. Gubbay, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7953-7957 (1992).
26. Tucker, P. K., Lee, B. K., Lundrigan, B. L. & Eicher, E. M. *Mammal. Genome* **3**, 254-261 (1992).
27. Altschul, S. F. & Ericson, B. W. *Bull. math. Biol.* **48**, 603-616 (1986).

ACKNOWLEDGEMENTS. We thank M. Potter (National Cancer Institute), E. Eicher (Jackson Laboratory), R. Sage (University of Missouri), B. Patterson and L. Heaney (Field Museum of Natural History), R. Baker (The Museum, Texas Tech University) for providing frozen organ tissue from wild mice; R. Lovell-Badge for sharing unpublished data; and B. Taylor, K. Miller, M. O'Keefe and S. Jansa for technical assistance. This publication was supported by a NSF grant to P.K.T.

§ To whom correspondence should be addressed.

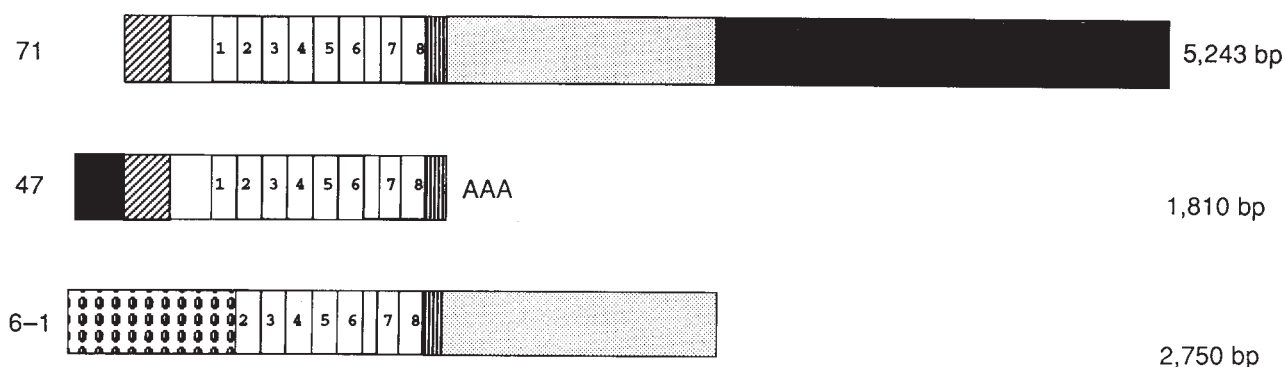
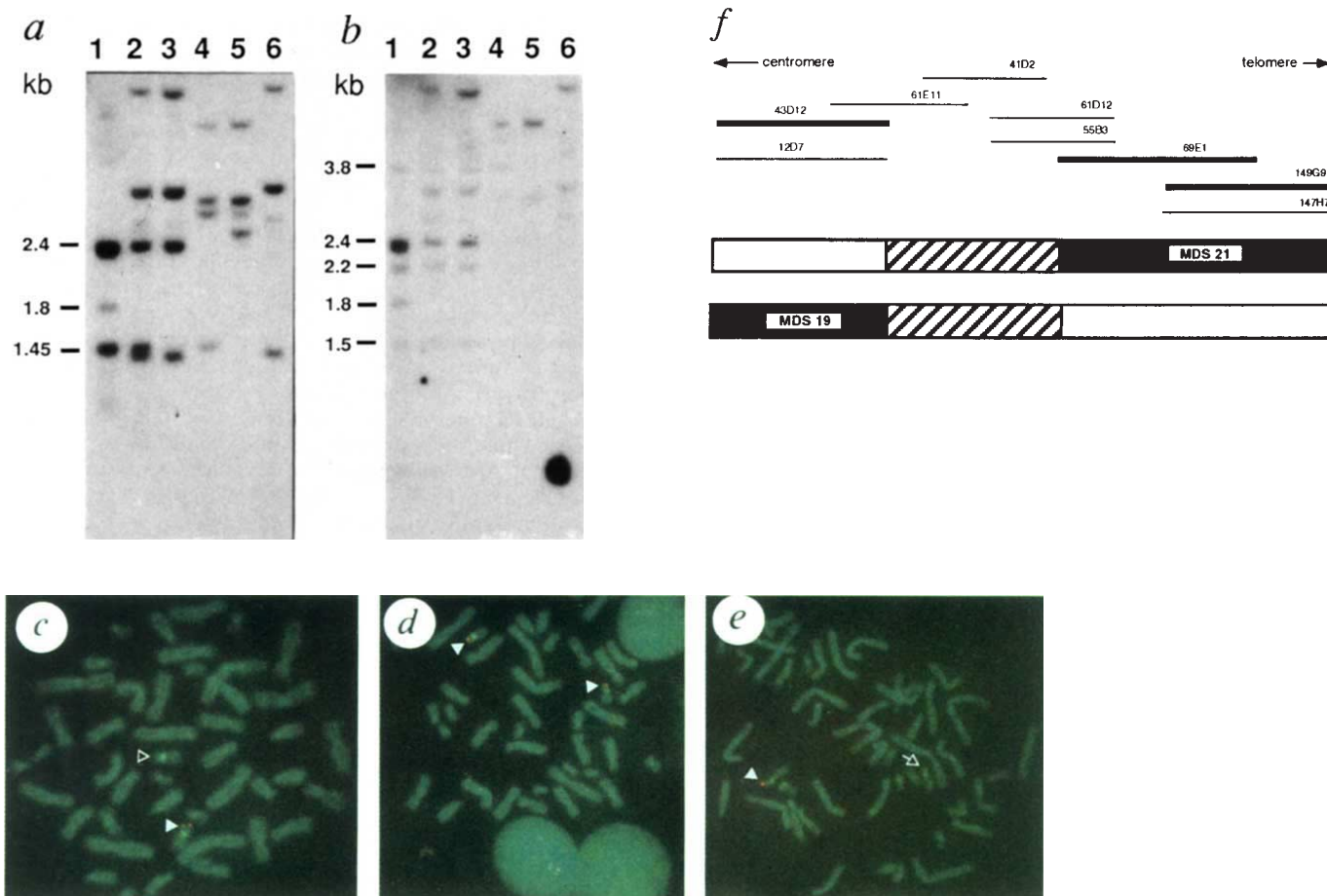


FIG. 1 Schematic representation of cDNA clones from chromosome 17. White areas correspond to open reading frames, amino-acid repeat units are denoted by numbers, the diagonally hatched areas at the 5' end of clones 71 and 47 represent identical 5' untranslated regions; dotted area represents a different 5' untranslated region in clone 6-1. Grey areas indicate identical 3' untranslated regions in clones 71 and 6-1, and black areas correspond to unique 5' or 3' sequences. Clones 47 and 71 were isolated from a human kidney cDNA library, and clone 6-1 was from a human fetal brain library (AAA indicates poly (A) tail). Comparison of DNA sequences using BLASTN²⁴ revealed several matches. Sequences in the 3' untranslated end of clone 71 identified two expressed sequence tags in Genbank²⁵, HUMXT00616 and HUMOS14G06. A fourth clone (clone 39) from the human kidney library had a different polyadenylation site corresponding to position 3,480 of clone 71. Clones 47 and 39 may account for the 2.2-kb and the 4.1-kb bands on the northern blot. A fifth clone (clone 8-1) from the human fetal brain library was partially sequenced. It contained 5' and 3'

sequences, was unstable, and identified several expressed sequence tags in Genbank: HUMOC13A08, T03737, HUMXT00621, HUMXT01298 and T03590. The nucleotide sequence of the open reading frame did not match any entry in Genbank. Sequences are deposited in Genbank under the following accession numbers: clone 71, L13385; clone 47, L13386; clone 6-1, L13387; clone 8-1 (partial sequence), L13388.

METHODS. A 1.8-kb *Eco*RI fragment from the clone that mapped to chromosome 2 was used to screen a λ -ZAP human fetal brain cDNA library (Stratagene) and a λ gt10 human kidney cDNA library (gift from G. Bell). Purified phages were excised *in vivo* from the λ ZAP and *Eco*RI fragments from the λ gt10 library were subcloned into Bluescript II KS+. For nucleotide sequencing an Applied Biosystems 373 fluorescent sequencer or manually Sequenase sequencing kit (USB, Cleveland) was used. Sequence comparisons were done with the University of Wisconsin GCG sequence analysis software.



the 5' region of the gene. In contrast hybrid JRL2 contains a maternal 17 homologue from patient MDS-21 (ref. 11) and shows a partial deletion involving only the 3' region of the gene. Patient MDS-21 is one of three MDS patients in whom all previous probes have failed to detect DNA alterations¹¹. The Southern blot analysis therefore indicates the presence of non-overlapping deletions of either the 5' or the 3' regions of the *LIS-1* gene in these two patients. Identical results were obtained by quantitative PCR (data not shown) and fluorescence *in situ* hybridization (FISH) (Fig. 2c-e).

By FISH analysis, patient MDS-19 showed hybridization to only one chromosome-17 homologue using cosmids from the 5' region of the gene (Fig. 2c), but both homologues hybridized with cosmids from the 3' end. Patient MDS-21 showed the opposite pattern: positive hybridization to both chromosome-17 homologues with cosmids from the 5' region of the gene (Fig. 2d), but to only one homologue with cosmids from the 3' end (Fig. 2e). These data confirm that the two patients have partial deletions at opposite ends of the *LIS-1* gene. Based on the FISH results and previous mapping data from these and other patients, the 5' end of the *LIS-1* gene is telomeric (summarized in Fig. 2f). Since these two patients have lissencephaly and the facial dysmorphism of MDS¹¹, and their non-overlapping deletions disrupt opposite ends of this cDNA, it seems that mutations of *LIS-1* can produce the complete MDS phenotype.

Northern blot analysis revealed a complex pattern of expression in human tissue (Fig. 3a), with transcripts of at least four different sizes (2.2–7.5 kb). Expression was detected in all tissues tested, but was most pronounced in brain, heart and skeletal muscle. Northern blot analysis of mouse tissues (Fig. 3b) revealed a major prominent band of 7.5 kb, which was most highly expressed in brain.

The deduced amino-acid sequence indicated that the *LIS-1* protein contains eight so-called 'WD40' repeats of the type found in G protein β -subunits⁷ (Fig. 4a, b). As the protein contains eight repeats and these repeats span most of the amino-acid

sequences, *LIS-1* seems to be most closely related to β -subunits of G proteins, although it is quite unlike any previously isolated β -subunits. The overall structure is most similar to the yeast β -subunit STE4 protein, because the amino terminus contains sequence unrelated to the repeat, and a hinge-like region between repeats five and six (in STE4 the hinge lies between repeats six and seven). Functional regions in the STE4 protein have been mapped by mutational analysis¹³. If *LIS-1* is a β -subunit of a G protein, then its relatively low sequence similarity (55%) to human β -transducin may indicate specificity for other components of the complex on the level of the β -subunit, similar to the specificity found in α - and γ -subunits^{14–18}.

G proteins participate in signal transduction in pyramidal neurons¹⁹, which are involved in the neuronal migration abnormalities in lissencephaly patients, as are most other neuronal types. Additionally, growth cone collapse, a normal process in neuronal migration, seems to be mediated by G proteins²⁰. Furthermore, the Alzheimer's amyloid protein precursor forms a complex with brain GTP-binding protein G₀, indicating that this complex may be involved in a neurodegenerative process²¹.

Although *LIS-1* may act in a signal transduction pathway crucial for brain development, it could be a member of the larger and functionally more diverse family of proteins that share the WD40 repeat unit. The β -transducin-like amino-acid repeat unit in *LIS-1* exists in several other proteins with unrelated functions, including proteins with eight repeat units such as slime mould AAC3, a developmentally regulated protein of unknown function, and yeast CDC4, which is essential for initiation of DNA replication and for separation of the spindle pole bodies during formation of the poles of the mitotic spindle. The *Drosophila* neurogenesis protein Groucho contains five of these repeats; yeast TUP1, which is a transcriptional repressor containing five repeats²² scored highest in a BLAST protein search.

Our results strongly suggest that *LIS-1* is the gene for lissencephaly and the associated facial dysmorphism of MDS which is located on chromosome 17. As deletions are hemizygous, half

FIG. 2 Detection of non-overlapping deletions in two patients with Miller-Dieker syndrome. a, b, Southern blot analysis of *Pst*I-digested genomic DNA from human control (lane 1); a mouse-human hybrid containing only human chromosome 17 (lane 2); mouse-human hybrid KCB4 containing the deleted chromosome 17 from patient MDS-19 (lane 3); hamster-human hybrid JRL2 containing the maternal homologue of chromosome 17 from patient MDS-21 (lane 4); control RJK88 hamster DNA (lane 5); and control CL1D mouse DNA (lane 6). In a, the probe was clone 8-1 and three human-specific bands are seen. The 1.8-kb band does not map to chromosome 17, as it is absent in the 17-only hybrid in lane 2. The 2.4-kb band is present in patient MDS-21 (lane 3) but is deleted in MDS-19 (lane 4). The 1.45-kb band, which migrates close to a slightly smaller mouse fragment (see doublet in lane 2 containing mouse and human chromosome 17), is present in MDS-19 (lane 4), but absent in MDS-21 (lane 3). In b, the probe was clone 6.1 and there are five human-specific bands (lane 1) but one of these (1.8 kb) does not map to chromosome 17 (lane 2). All four of the chromosome-17-specific bands are present in MDS-19 (lane 3), but only one of these (3.8 kb) is retained in patient MDS-21 (lane 4). A second hybrid clone of MDS-21 containing the maternal homologue confirmed this deletion; a clone containing the paternal homologue was normal. As the patients have different fragments that are deleted and have the single 3.8-kb fragment in common, their deletions are non-overlapping. c–e, Fluorescence *in situ* hybridization (FISH) analysis of metaphase chromosomes from patients MDS-19 and MDS-21. For these experiments, cosmids were isolated by screening a chromosome-17-specific cosmid library from the Los Alamos National Library. For all figures, a chromosome-17 α -satellite probe was used to identify the two chromosome 17 homologues unambiguously and these appear as green (biotin-FITC) centromeric signals. Cosmid probes appear as red signals (digoxigenin-rhodamine) at distal 17p. c, Patient MDS-19 shows a positive signal for the 5' region cosmids on one homologue (filled arrowhead) and a deletion on the other homologue (open arrowhead). d, Patient MDS-21 gives positive signals for the 5' region cosmids on both homo-

logues (filled arrowheads). e, Patient MDS-21 shows positive hybridization for the 3' region cosmids on one homologue (filled arrowhead), but is deleted on the other homologue (open arrowhead). f, Schematic representation of the deletions detected in patients MDS-19 and MDS-21 in reference to a cosmid contig constructed from 9 cosmids isolated from the chromosome-17-specific library. Cosmids 69E1 and 149G9 were used in FISH experiments to represent the 5' portion of the *LIS-1* gene, and showed a deletion in MDS-19 (open rectangle) but were present in MDS-21 (filled rectangle). Cosmid 43D12 was used in FISH experiments to represent the 3' portion of the gene; it was deleted in patient MDS-21 (open rectangle) but was present in patient MDS-19 (filled rectangle). The breakpoints of the deletions for each of the two patients lies within the hatched area, although Southern blot data indicates that they do not overlap one another.

METHODS. Somatic hybrid cell lines were constructed as described²⁶ by fusion of patient lymphoblasts to either mouse or hamster cells, followed by selection in hypoxanthine/aminopterin/thymidine medium. Hybrid clones containing only one chromosome-17 homologue were identified by PCR screening of YNZ22 (D17S5) alleles and confirmed by cytogenetic analysis. Standard procedures were used for digestion of DNA with *Pst*I and Southern blot analysis²⁷; FISH has been described^{28,29}. Metaphase chromosomes were prepared from lymphoblastoid cell lines from patients MDS-19 and MDS-21. One cosmid clone representing the 3' portion of *LIS-1* and two cosmids representing the 5' portion were co-hybridized with chromosome 17 α -satellite (D17Z1). Cosmids were labelled by nick-translation with digoxigenin-11-dUTP and detected with anti-digoxigenin [Fab]-rhodamine (Boehringer-Mannheim) without amplification. D17Z1 was labelled by nick-translation with biotin-11-dUTP and detected with FITC-avidin (Vector) without amplification. Slides were counterstained with DAPI (4',6-diamidino 2-phenylindole dihydrochloride) and visualized through a triple band-pass filter set (Zeiss) which allowed simultaneous detection of DAPI, FITC and rhodamine. At least 20 cells were scored for each cosmid set on each patient.

the normal dosage of the protein product is apparently insufficient for normal development. It may be that improper proportions of β - and γ - subunits of a G protein disturb formation of the normal protein complex, as in haemoglobin H disease, which is caused by an imbalance in the ratio of α - to β -globin²³. Further experiments are required to characterize other genes in

this apparent family, in particular the homologue on chromosome 2, which is almost identical to the chromosome-17 *LIS-1* cDNA. Genotype/phenotype studies could eventually explain the variability of phenotypic expression among MDS and isolated lissencephaly patients. Investigation of *LIS-1* expression during mouse CNS development may help to elucidate the func-

FIG. 3 *LIS-1* expression in different tissues. *a*, Northern blot of 2 μ g of poly(A)⁺ RNA from human tissues. Comparable loading of RNA in each lane was verified by reprobing the blot with a β -actin probe (data not shown). *b*, Northern blot of total RNA (15 μ g) from mouse tissues; loading RNA in each lane is variable. Note the strong expression of *LIS-1* in brain. METHODS. The northern human blot was purchased from Clontech. RNA was purified from mouse tissues by tissue homogenization in guanidine isothiocyanate and separated on a 1.5% formaldehyde-agarose gel, transferred to a nylon membrane and probed with a random-primer-labelled probe representing the *LIS-1* coding region and a β -actin probe. Filters were washed at high stringency (65 °C, 0.1 $\times$ SSC) and exposed to Kodak X-ray film overnight.

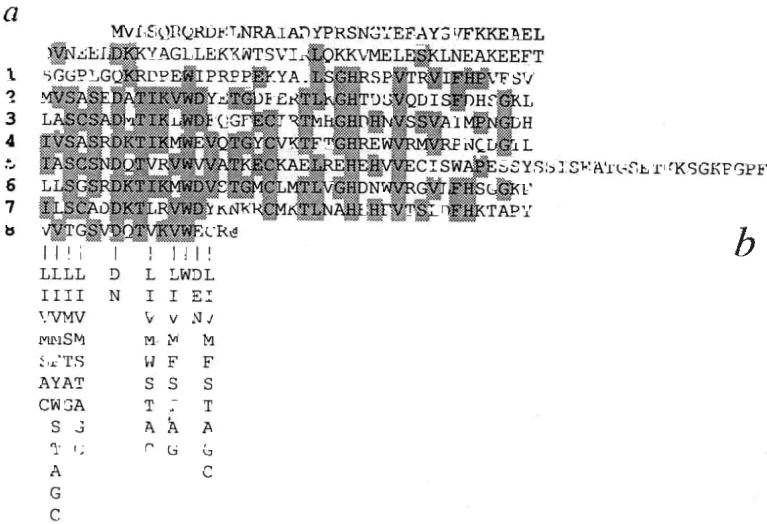
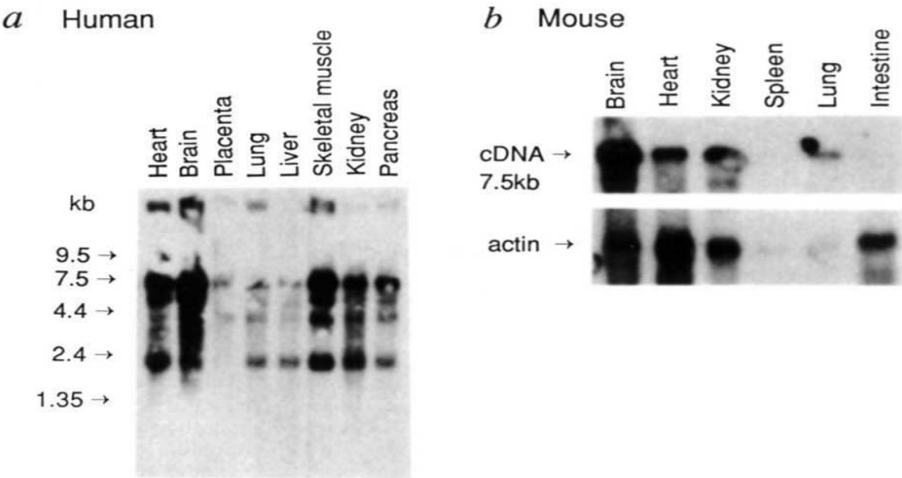
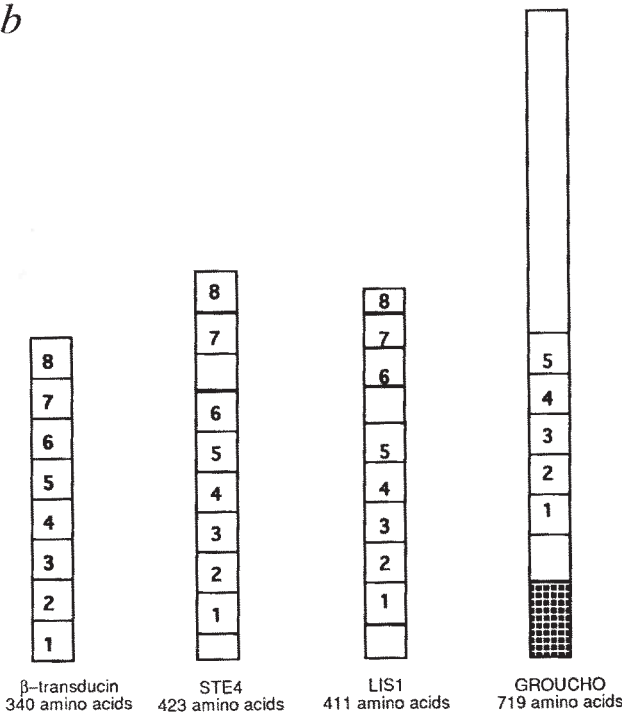


FIG. 4 Repeat domains in the deduced amino-acid sequence of *LIS-1*. *a*, Alignment of repeat regions (411 amino acid residues). Numbers on the left indicate the numbers of the repeat unit. Repeats were identified in the PROSITE database (updated June 1992) as β -transducin family Trp-Asp repeat signatures. The pattern included in the PROSITE database does not detect all occurrences of the Trp-Asp or WD40 domains because some repeats are less conserved than others, but it does detect all proteins containing this repeat. Shaded background shows identities in amino acids between domains; two gaps were included for better alignment. The PROSITE pattern is aligned below the amino-acid sequence. Sequenced clones 71 and 47 contain the entire *LIS-1* open reading frame. Clone 6-1, which has a different 5' end, starts with the first methionine in repeat 2 (residue 121). *b*, Schematic representation of repeat domains in four members of the β -transducin family. The β -transducin-like repeat elements are shown as white boxes marked with the repeat number. The black-checked box in the Groucho protein denotes a region of charged residues.



tion of this gene. Identification of the *LIS-1* gene should therefore prove useful in the study of neuronal migration in normal development as well as in this unique human disorder. □

Received 2 April; accepted 27 May 1993.

1. Aicardi, J. *Int. Pediatr.* **4**, 118–126 (1989).
2. Dobyns, W. B. *Neurol. Clin.* **7**, 89–105 (1989).
3. Barth, P. G. *Can. J. Neurol. Sci.* **14**, 1–16 (1987).
4. Miller, J. Q. *Neurology* **13**, 841–850 (1963).
5. Dieker, H. et al. *The Lissencephaly Syndrome. The Clinical Delineation of Birth Defects II: Malformation Syndromes 1–64* (National Foundation March of Dimes, New York, 1969).
6. Ledbetter, S. A., Kuwano, A., Dobyns, W. B. & Ledbetter, D. H. *Am. J. hum. Genet.* **50**, 182–189 (1992).
7. Duronio, R. J., Gordon, J. I. & Boguski, M. S. *Proteins* **13**, 41–56 (1992).
8. Lee, C. C. et al. *Science* **239**, 1288–1291 (1988).
9. Guillemot, F., Billault, A. & Auffray, C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4594–4598 (1989).
10. Guzzetta, V. et al. *Genomics* **13**, 551–559 (1992).
11. Dobyns, W. B., Curry, C. J. R., Hoyme, H. E., Turlington, L. & Ledbetter, D. H. *Am. J. hum. Genet.* **48**, 584–594 (1991).
12. Ledbetter, D. H. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5136–5140 (1989).
13. Leberer, E., Dignard, D., Hougau, L., Thomas, D. Y. & Whiteway, M. *EMBO J.* **11**, 4805–4813 (1992).
14. Birnbaumer, L. *Cell* **71**, 1069–1072 (1992).
15. Simon, M. I., Strathmann, M. P. & Gautam, N. *Science* **252**, 802–808 (1991).
16. Gilman, A. G. *Rev. Biochem.* **56**, 615–649 (1987).

17. Kleuss, C., Scherubel, H., Hescheler, J., Schultz, G. & Wittig, B. *Science* **259**, 832–834 (1993).
18. Iniguez-Lluhi, J. A., Simon, M. I., Robishaw, J. D. & Gilman, A. G. *J. bio. Chem.* **267**, 23409–23417 (1992).
19. Bourne, H. R. & Nicholl, R. *Cell* **72**, 65–76 (1993).
20. Igarashi, M., Strittmatter, S. M., Vartanian, T. & Fishman, M. C. *Science* **259**, 77–79 (1993).
21. Nishimoto, I. et al. *Nature* **362**, 75–79 (1993).
22. Keleher, C. A., Redd, M. J., Schultz, J., Carlson, M. & Johnson, A. D. *Cell* **68**, 709–719 (1992).
23. Orlin, S. H. & Nathan, D. G. *New Engl. J. Med.* **295**, 710–714 (1976).
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403–410 (1990).
25. Bilofsky, H. S. & Burks, C. *Nucleic Acids Res.* **16**, 1861–1864 (1988).
26. vanTuinen, P., Rich, D. C., Summers, K. M. & Ledbetter, D. H. *Genomics* **1**, 374–381 (1987).
27. Sambrook, S., Fritsch, E. I. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
28. Kuwano, A., Ledbetter, S. A., Dobyns, W. B., Emanuel, B. S. & Ledbetter, D. H. *Am. J. hum. Genet.* **49**, 707–714 (1991).
29. Kuwano, A. et al. *Hum. molec. Genet.* **1**, 417–425 (1992).

ACKNOWLEDGEMENTS. We thank M. J. McKinney for technical assistance, and L. Birnbaumer, A. Gilman, R. Smith and B. Rossiter for critical review of the manuscript. O.R. was supported by a Human Frontier Science Program Organization long-term fellowship and is currently supported by a Fogarty fellowship. This work was supported by the Baylor Human Genome Center and by grants from the Department of Energy, the NIH (to D.H.L. and W.B.D.) and the Deutsche Forschungsgemeinschaft (M.W.); F.F. is supported by a Duchenne muscular dystrophy fellowship. C.T.C. is an Investigator with the Howard Hughes Medical Institute.

Spectral motion produces an auditory after-effect

Z. J. Shu^{*†}, N. V. Swindale^{*} & M. S. Cynader^{*}

^{*}Department of Ophthalmology, University of British Columbia, 2550 Willow Street, Vancouver, British Columbia V5Z 3N9, Canada

[†]Department of Electrical Engineering, University of British Columbia, 2356 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada

DISTORTIONS of perception following prolonged exposure to an unvarying sensory stimulus have been observed since at least the third century BC¹. The motion after-effect is a familiar experience² in which, after a few minutes of viewing objects moving in a single direction, a stationary object appears to move in the opposite direction. Similar after-effects have been observed for many visual stimuli, including tilted lines, colours, stereoscopic depth, curvature, spatial frequency, contrast, rotation and motion in depth^{3–9}. In contrast to the rich variety of visual after-effects reported since the 1960s, reports of analogous auditory adaptation effects only appeared in the 1970s^{10–12}, but have continued since then^{13,14}. Some effects of sound source spatial movement perception after adaptation to a spatially moving sound source have been reported¹⁵. Here we report an auditory perceptual after-effect analogous to the visual motion after-effect, which is caused by adaptation to auditory spectral (frequency) motion. After a few minutes of listening to a simple spectral pattern moving upwards or downwards in frequency space, the same pattern sounds as though it is drifting in the opposite direction when it is stationary. The effect shows binaural transfer, implying that it is generated at the level after binaural interaction. After-effects produced by the motion of spectral peaks are independent of those produced by spectral notches, suggesting separate processing channels for spectral peaks and notches.

The stimulus had a broad band, flat, noise spectrum on which was superimposed either a peak of narrow band noise, centred on a given frequency, or a narrow band notch, likewise centred on a given frequency (Fig. 1). The centre frequency of the peak or notch moved either upwards or downwards in frequency space, at a constant velocity, and with a repetitive sawtooth pattern. A two-alternative forced choice procedure was used to measure the after-effect of adaptation to this spectral motion. After 2–3 minutes of exposure to an adapting stimulus, the listener was asked to judge the direction of motion (up or down in frequency) of a briefly presented (0.5 s or less) test stimulus. Each test stimulus could have one of seven different velocities,

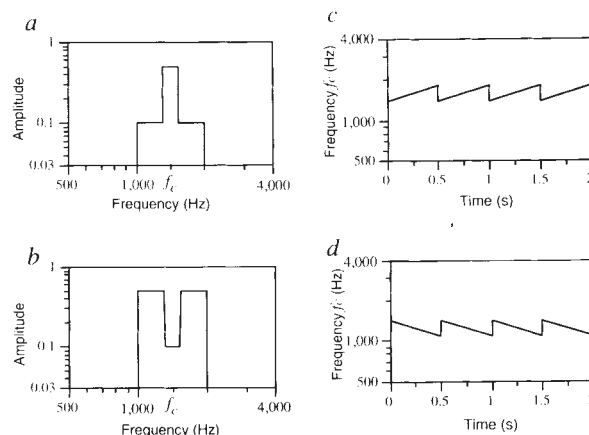


FIG. 1 Instantaneous amplitude spectra of the moving peak (a) and notch (b) stimuli, and the time functions of the central frequency of the peak or notch (c and d). The stimulus had a broad band, flat, noise spectrum on which was superimposed either a peak of narrow band noise, centred on frequency f_c , or a narrow band notch in the spectrum, also centred on f_c . f_c moved either upwards (c) or downwards (d) in frequency, at a constant velocity in log frequency space, and in a repetitive sawtooth pattern. The initial value of f_c was logarithmically centred in the frequency band of 1 to 2 kHz. The peak or notch was about 200 Hz wide and more than 15 dB above or below the baseline, which was kept constant during stimulus presentation. The velocities used in the briefly presented test stimuli were always slow enough that f_c did not cross the boundaries of the 1–2 kHz band during the test stimulus interval.

METHODS. The sound signals were computed digitally by two Motorola DSP56001 microprocessors in a Macintosh II computer. The sampling rate was 14,700 Hz, and the additive synthesis method was used. Sixty-four sinusoidal components with randomized initial phases were synthesized using the table look-up method with fractional phase steps and a table size of 1,024. Analogue voltages generated by the DSP boards were passed to a stereo amplifier (NAD Electronics, London) and presented to the subject by electrostatic headphones (Stax Industries, Tokyo).

chosen at random on each trial. Following each test, the adapting stimulus was presented again for 2 s to maintain the adaptation. A total of about 100 test and reinforcing presentations was given on each test of adaptation. The range of test velocities was always centred on zero and was chosen so that the extremes of the range were nearly always correctly identified by the subject

as moving up or down. Probit analysis¹⁶ was used to estimate the 50% response rate on the resulting psychometric function. This is the stimulus velocity which sounded stationary to the listener (unchanging in pitch). A significant change in this subjective mean following adaptation indicates the presence of an after-effect.

In the first three experiments we examined the after-effect using either similar or different adapting and test stimuli. In experiment 1 we measured the after-effect obtained when the adapting and test stimuli were both moving peak stimuli (Fig. 1a). Different adapting velocities were used to examine their effects on the size of the after-effect. All subjects showed significant changes in the estimated subjective mean velocities after adaptation as compared with that of the control experiment (Fig. 2a). When the adapting velocity was positive, that is, upwards in frequency, the subjective mean velocity changed in a positive direction. Thus a positive velocity was judged by listeners as stationary and correspondingly a stationary velocity was heard as moving downwards. At negative adapting velocities, a moving peak stimulus with a negative velocity was also judged as stationary. The after-effects were maximum at adapting velocities of around ± 0.5 octaves per s.

For experiment 2, a different spectral pattern was used. At any given instant during presentation, both the adapting and the test stimuli had a notch in an otherwise flat spectrum over the same frequency band as that used above (Fig. 1b). The bandwidth of the notch was 200 Hz, the same as that of the peak in the first experiment. The centre frequency of the notch moved up or down in frequency space and, as before, the task of the listener was to judge the direction of motion of the notch in the test stimuli. All subjects showed significant after-effects following adaptation for adapting velocities around ± 0.5 octaves per s. In Fig. 2b the size of the after-effect is plotted as a function of the adapting velocity. The results were similar to those of experiment 1: after adapting to upward motion of a moving notch, a stationary notch appeared to move downwards, and vice versa.

To test whether adaptation to moving peaks and notches involved the same processing channels (experiment 3), we adapted first to a moving peak and then tested with a notch, and vice versa (cross-adaptation). The flat background level of the adapting stimulus and that of the test stimulus were kept the same. Figure 2c and d show the effect of adapting velocity on the subjective mean velocity of the test stimulus. Although there was some evidence for a weak effect of adaptation to a moving peak when tested with a notch stimulus, cross-adaptation generally produced effects that were weak or arguably non-existent. This suggests that after-effects produced by moving peaks and notches occur in different sensory channels.

In the visual system, after-effects thought to have a central locus typically show transfer between the two eyes (such as the motion after-effect), whereas after-effects that involve peripheral mechanisms (such as coloured after-images) do not transfer. To test whether the auditory after-effect is monaural or involves binaural pathways we adapted one ear and tested with the other ear. Figure 3 shows the size of interaural transfer of the after-effect as a function of adapting velocity. The after-effects for interaural transfer and same-ear adaptation were similar, suggesting that they are generated after the site of binaural interactions in the nervous system, and ruling out peripheral mechanisms for the after-effect.

These results suggest that the auditory system has a previously unsuspected sensitivity to the motion of peaks and troughs in the spectrum of complex sounds, extending the analogy between position on the basilar membrane of the cochlea and the position on the retina. Sensitivity to shape cues in the sound spectrum is an important factor in monaural sound localization¹⁷⁻¹⁹. The filtering effects of the head, pinna and the ear canal are substantial and complex, and result in changes in the spectral content of incoming sounds that are dependent on the incidence angle of the sound wavefront relative to the head. Such spectral cues

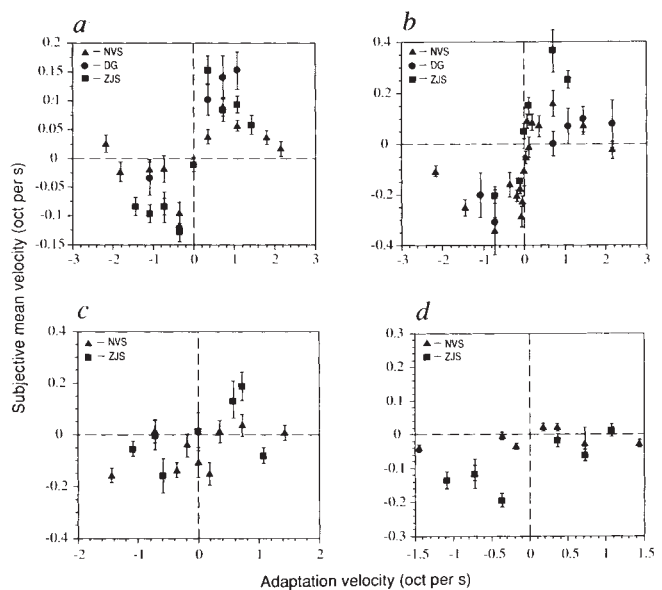


FIG. 2 Size of the monaural adaptation after-effect (in terms of the subjective mean velocity at which the stimulus sounded stationary) as a function of the velocity of the central frequency of the peak or notch in the adapting stimulus. Error bars indicate one s.d. in the mean, as estimated by the Probit procedure. Three listeners (N.V.S., D.G. and Z.J.S.) conducted these experiments. a, The listeners adapted to, and were tested with, moving peak stimuli; b, both the adapting and test stimuli were moving notches; c, the adapting stimulus was a moving peak whereas the test stimuli were moving notches; d, the adapting stimulus was a moving notch whereas the test stimuli were moving peaks. See text for details of the experiments, and Fig. 1 for the stimuli.

for monaural localization are thought to be important in locating the elevation of sound sources, which cannot be done using purely binaural cues, and may rely on detection and discrimination of spectral peaks and notches²⁰.

If, as seems likely, the auditory motion after-effect reported here is a consequence of changes of activity within populations of neurons selectively sensitive to spectral motion, the results of experiment 3 suggest that different neuronal populations encode the motion of peaks and notches in auditory stimuli. In addition, if sound frequency is analogous to space in visual perception, our finding of an after-effect for motion peaks or notches in

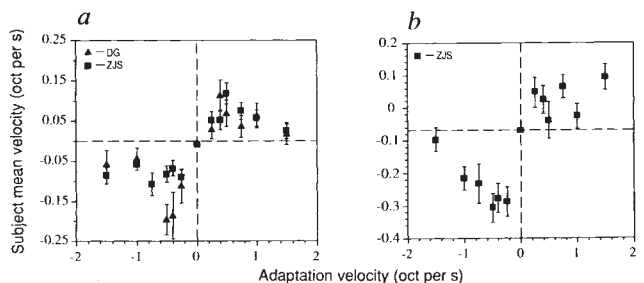


FIG. 3 Size of interaural transfer of the after-effect (in terms of the subjective mean velocity at which the stimulus sounded stationary) as a function of the velocity of the central frequency of the peak or notch in the adapting stimulus. In this case one ear was adapted and the other ear tested. Crosstalk between the two ears in our system was carefully checked to ensure that stimulation in the adapting ear was inaudible in the test ear. Error bars indicate one s.d. in the mean, as estimated by the Probit procedure. a, The adapting and test stimuli were both moving peaks; b, the adapting and test stimuli were both moving notches.

frequency space is simply an extension of the well known visual motion after-effect to another sensory domain. By analogy there may be neurons in the auditory pathways that are sensitive to the motion of spectral patterns, implying the existence of the auditory equivalent of dark and light bar detectors in the visual system. The results of interaural transfer suggest that these detectors are located at a level after binaural interaction. Although there is evidence for neurons sensitive to the motion of frequency peaks²¹ there are as yet no reports of neurons sensitive to moving spectral notches, or even stationary gaps in the spectrum. There is evidence that spatial auditory source movement can be sensitively signalled by binaural temporal disparities in frequency-modulated signals¹³, however this is likely to involve mechanisms other than those that detect changing frequency²². Our results suggest that further attempts to define such channels with neurophysiological and psychophysical methods would be worthwhile. □

Received 29 March; accepted 24 May 1993.

1. Aristotle *Parva Naturalia*.
2. Wohlgemuth, A. *Brit. J. Psychol.* **1**, 1–117 (1911).
3. Barlow, H. B. & Hill, R. M. *Nature* **200**, 1345–1347 (1963).
4. Sekuler, R. W. & Ganz, L. *Science* **139**, 419–420 (1963).
5. Anstis, S. M. *Science* **155**, 710–712 (1967).
6. Mayhew, J. E. W. & Anstis, S. M. *Percept. Psychophys.* **12**, 77–85 (1972).
7. Walker, J. T. *Psychonom. Sci.* **28**, 333–335 (1972).
8. Potts, M. J. & Harris, J. P. *Vision Res.* **15**, 1225–1230 (1975).
9. Favreau, O. E. *Vision Res.* **16**, 181–186 (1976).
10. Kay, R. H. & Matthews, D. R. *J. Physiol., Lond.* **225**, 657–677 (1972).
11. Regan, D. & Tansley, B. W. *J. Acoust. Soc. Am.* **65**, 1249–1257 (1979).
12. Gardner, R. B. & Wilson, J. P. *J. Acoust. Soc. Am.* **66**, 704–709 (1979).
13. Kay, R. H. *Physiol. Rev.* **62**, 894–975 (1982).
14. Wilson, J. P., Crampin, E. J. & Cann, N. *Brit. J. Audiol.* **26**, 188 (1992).
15. Grantham, D. W. *Percept. Psychophys.* **45**, 129–136 (1989).
16. Finney, D. J. *Probit Analysis* 3rd edn (Cambridge Univ. Press, 1971).
17. Blauert, J. *Spatial Hearing* (MIT Press, Cambridge, MA, 1983).
18. Bloom, P. J. *J. Audio. Eng. Soc.* **25**, 560–565 (1977).
19. Butler, R. A. & Helwig, C. C. *Am. J. Otolaryngol.* **4**, 165–173 (1983).
20. Moore, B. C. J., Oldfield, S. R. & Dooley, G. J. *J. Acoust. Soc. Am.* **85**, 820–836 (1989).
21. Mendelson, J. R. & Cynader, M. S. *Brain Res.* **327**, 331–335 (1985).
22. Green, G. G. R., Heffer, R. J. S. & Ross, D. A. *J. Physiol., Lond.* **260**, 49P (1976).

ACKNOWLEDGEMENTS. We thank D. Glaschi for participating in the experiments as a subject, R. M. Douglas for helping with the computer system used in our experiments and C. A. Laszlo for his advice and encouragement to Z.J.S. This research was supported by a MRC (Canada) grant to M.S.C. and by the Institute for Robotics and Intelligent Systems (Canada).

Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state

Patricio T. Huerta & John E. Lisman

Department of Biology and Center for Complex Systems, Brandeis University, Waltham, Massachusetts 02254–9110, USA

BRAIN cholinergic neurons are critical for memory function^{1,2} and their loss may contribute to memory impairment in Alzheimer's disease³. One role of cholinergic neurons is to elicit an oscillatory activity called theta rhythm⁴ in the hippocampus, a brain region involved in memory processing⁵. Theta rhythm occurs during periods of learning^{6,7}, but its effect on the synaptic plasticity that underlies learning remains unclear. We have studied synaptic plasticity in hippocampal slices during theta-frequency oscillations induced by a cholinergic agonist^{8–10}. Here we report that during these oscillations, synapses are in a state of heightened plasticity and can be modified by what would otherwise be ineffective stimulation. This heightened plasticity is sensitive to the timing of incoming stimuli with respect to the oscillatory activity. The results suggest that cholinergic systems may affect memory formation through the induction of an oscillatory state in which the requirements for synaptic plasticity are dramatically altered.

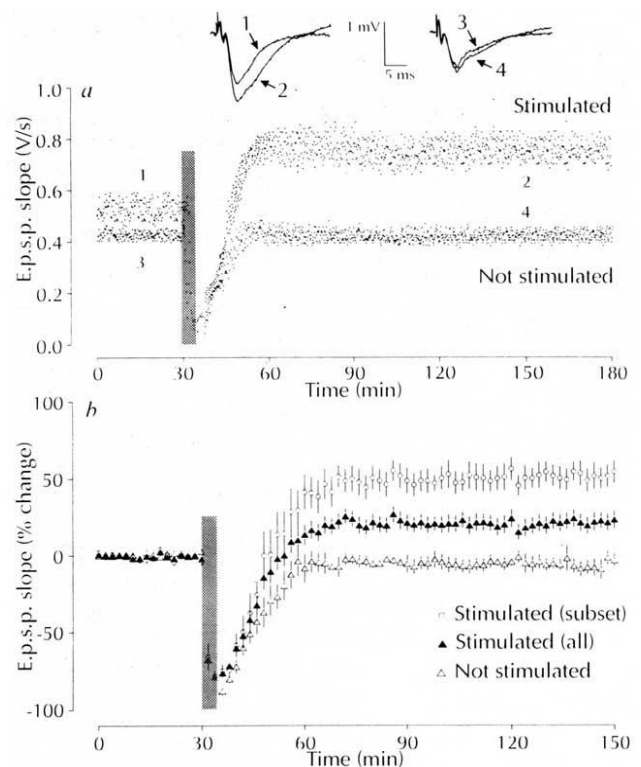
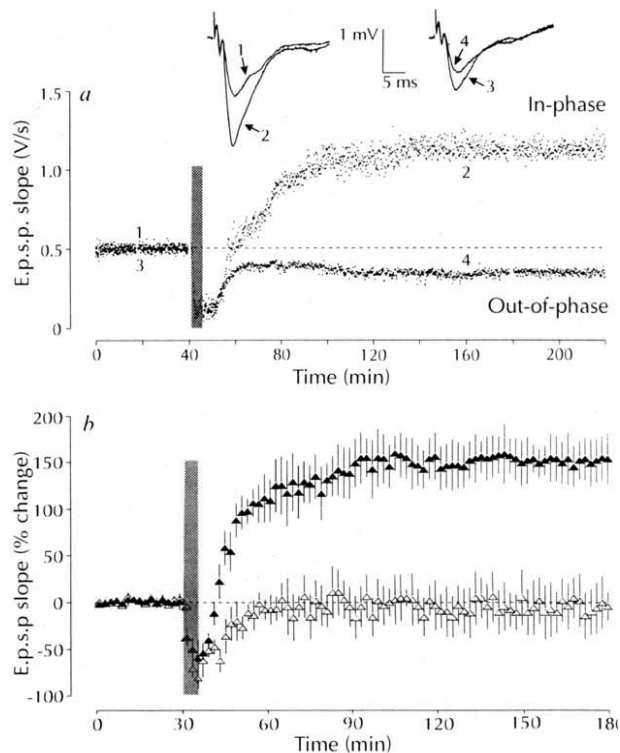


FIG. 1 Cholinergically induced long-term enhancement of synaptic efficacy. *a*, Representative experiment showing the slope of the field e.p.s.p. against time. During bath application of CCh (50 μ M for 5 min, hatched rectangle), one pathway was stimulated and another pathway was not. Theta-frequency oscillations (mean amplitude, 280 μ V) were present during CCh, but stimulation was not synchronized with the oscillations. Inset, field e.p.s.ps (average of 6 traces) taken at the times indicated by the numbers in the graph. *b*, Pooled data ($n=27$) showing the average per cent change of the field e.p.s.p. in two pathways; synapses that were stimulated ($\blacktriangle$) and synapses that were not stimulated ($\triangle$) during CCh. Also shown are results from the stimulated pathway in the subset of experiments ($n=10$) that showed large (>100 μ V) theta-frequency oscillations during CCh ($\circ$). Symbols represent mean change ($\pm$ s.e.m.) plotted at 2-min intervals (intervening data have been omitted for clarity). Similar results were obtained for the population spike amplitude concurrently recorded in stratum pyramidale in a subset of experiments: neurons that were stimulated during CCh were persistently enhanced ($49 \pm 4\%$, $n=23$, $P<0.001$, not shown).

METHODS. Transverse hippocampal slices (400 μ m) were obtained from 100–250 g (2–8 weeks old) male Long-Evans rats²⁹ and maintained in a submerged chamber with continuous perfusion of oxygenated solution (in mM): NaCl, 124; NaH_2CO_3 , 26; D-glucose , 10; KCl, 5; CaCl_2 , 2; MgSO_4 , 2; NaH_2PO_4 , 1.2 (pH 7.4) at 30–32 $^\circ\text{C}$. An extracellular recording microelectrode (2 M NaCl, 1–3 M Ω) was placed in stratum radiatum, 150–250 μ m from stratum pyramidale. Recordings were made with an Axoclamp 2A amplifier. Evoked responses were sampled at 10 KHz and analysed on-line to determine the field e.p.s.p. initial slope. Two bipolar Pt/Ir microelectrodes were placed onto Schaffer collateral axons (at opposite sides of the recording microelectrode) and delivered current stimuli (0.1–0.5 mA, 50 μ s) to elicit half-maximal response; stimuli were alternated so that a response was recorded every 5 seconds. To ensure independent stimulation of two sets of synapses, a paired-pulse facilitation paradigm (inter-stimulus interval of 50 ms) was used. Summary graphs were constructed by: (1) normalizing each experiment, by expressing all values as percentages of the mean value before CCh application, (2) aligning them with respect to the onset of CCh, and (3) averaging the time-matched, normalized data across experiments. The per cent changes quoted in the text (mean change $\pm$ s.e.m.) were obtained from the summary data by comparing the distributions of 18 averaged values during two 3-min intervals, one just before CCh application and the other 2 h after CCh application. Statistical significance, P , of this per cent change was computed using an unpaired Student t -test.

FIG. 3 Selective enhancement of synapses stimulated at positive peaks of the theta-frequency waves. **a**, Representative experiment showing the slope of the field e.p.s.p. against time. During CCh application (hatched rectangle) one pathway was stimulated in-phase and another pathway was alternately stimulated out-of-phase, at 10-s intervals for each pathway. Inset, field e.p.s.ps (average of 6 traces) taken at the times indicated by the numbers in the graph. **b**, Pooled data ($n=5$) showing the average per cent change of the field e.p.s.p. in the two pathways; in-phase inputs ($\blacktriangle$) showed stable long-term enhancement while out-of-phase inputs ($\triangle$) showed little change following CCh treatment. Symbols represent mean change ($\pm$ s.e.m.) plotted at 2-min intervals, for clarity. In other experiments, we found that giving the out-of-phase stimulation alone produced no change ($-2 \pm 4\%$, $n=4$, not shown). **METHODS.** The onset of theta-frequency waves was detected on a chart record of field potentials during CCh application. When large-amplitude waves ($>100 \mu\text{V}$) appeared, a software routine was invoked which first found the maximum and minimum values in an 8-s period. Thresholds for synaptic stimulation were set at 95% of the maximum or minimum voltages. The voltage was then sampled at 1 kHz and a single stimulatory pulse was given when voltage exceeded the appropriate threshold. The interval between stimulations of a given pathway was ~ 10 s and the interval between in-phase and out-of-phase stimulation was ~ 5 seconds. The entire period of 'phase' stimulation was ~ 5 min (30–40 stimuli per pathway).



Schaffer collateral/commissural inputs onto CA1 neurons were stimulated at 0.1 Hz with single shocks, each of which evoked a field excitatory postsynaptic potential (e.p.s.p.) (Fig. 1a, inset). Application of carbachol (CCh), a cholinergic agonist, reduced the size of the e.p.s.p.^{11,12}. Following CCh removal, however, the e.p.s.p. increased to a level substantially larger than before CCh application (Fig. 1; change was $22 \pm 6\%$ measured 2 hours after CCh removal, $n=27$, $P<0.005$). This enhancement lasted as long as the recordings were maintained (≤ 3.5 h). To determine whether the cholinergically induced enhancement required synaptic activation, a second set of CA1 synapses were not stimulated during CCh application. As shown in Fig. 1, synapses that were not stimulated during CCh were not modified (change was $-4 \pm 3\%$, $n=27$). This indicates that the enhancement requires presynaptic stimulation and is synapse-specific. As is well established, the same low-frequency stimulation (0.1 Hz) in the absence of CCh does not produce changes in the e.p.s.p. Thus in CCh, a long-term enhancement is induced by stimulation that would normally cause no modification in synaptic efficacy.

As mentioned above, CCh elicits an oscillatory extracellular potential in hippocampal slices^{8–10} in the same frequency range (4–9 Hz) as the acetylcholine-dependent theta rhythm recorded

*in vivo*⁴. Both oscillations have strong similarities¹³ and are correlated with rhythmic neuronal spiking^{13,14} (but see ref. 15). Theta-frequency waves (Fig. 2) were observed in $\sim 75\%$ of our experiments and only occurred during CCh application. Notably, in the experiments in which CCh elicited robust oscillations ($>100 \mu\text{V}$) there was a pronounced enhancement (Fig. 1b; change was $52 \pm 8\%$, $n=10$, $P<0.001$). Conversely, in the cases with no detectable oscillation or small oscillations, the enhancement was negligible (change was $3 \pm 1\%$, $n=17$, not shown). This suggests that the enhancement produced by CCh is not purely a biochemical process, but depends on the induction of an electrical oscillation.

To explore further the importance of the theta-frequency oscillations we synchronized the stimuli to a given phase of the waves. Figure 3 shows the outcome of experiments in which one input pathway was stimulated by single pulses given at positive peaks of the wave ('in-phase') while another pathway was stimulated in alternation at the negative valleys ('out-of-phase'), as detailed in Fig. 3 legend. In-phase stimulation greatly enhanced the e.p.s.p. (for example, Fig. 3a). In fact, the enhancement due to in-phase stimulation was considerably larger (Fig. 3b; change was $151 \pm 5\%$, $n=5$, $P<0.001$) than that produced by giving

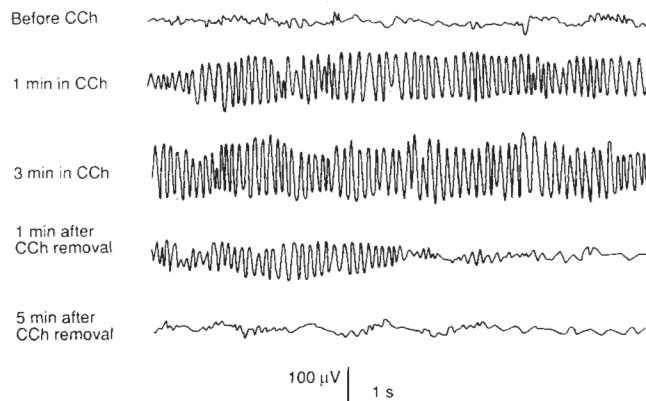


FIG. 2 Rhythmic oscillations of theta frequency are excited by CCh. Voltage traces represent activity before, during and after CCh application.

METHODS. The recording electrode placed in stratum radiatum was used to measure theta-frequency waves. These signals were filtered (1–25 Hz) and stored in a computer (sample rate of 1 kHz) and a chart recorder.

stimulation not synchronized with respect to the theta oscillations (that is, $52 \pm 8\%$) as in Fig. 1. In the second pathway, out-of-phase stimulation produced no enhancement (Fig. 3b; change was $-7 \pm 2\%$, $n = 5$), and in 2 of 5 cases there was a clear decrease (for example, Fig. 3a). In the experiments of Fig. 3, the two pathways were stimulated alternately, but similar results were obtained in experiments using only one type of stimulation. These experiments demonstrate that synchronizing stimulation to the phase of theta waves dramatically affects the amount of long-term enhancement produced. They also strongly suggest that the electrical consequences of cholinergic action are responsible for the heightened plasticity.

These results add a new perspective on the role of oscillatory states in the central nervous system. Previous studies have focused on the computational importance of brain oscillations^{16,17}. Our results, and related *in vivo* work in the dentate gyrus¹⁸, indicate that oscillations also have consequences for synaptic plasticity. Specifically, in the oscillatory state, the phase relation between incoming stimuli and the oscillation strongly affects the resulting change in synaptic efficacy. Thus, if incoming activity to the hippocampus is at theta frequency^{7,14}, only information of the appropriate phase will be stored. This demonstration of the importance of theta frequency is consistent with previous work showing that hippocampal plasticity is especially tuned to theta-frequency inputs even in the absence of intrinsic rhythmic activity¹⁹⁻²².

Studies on hippocampal and neocortical neurons have shown that acetylcholine affects a variety of biochemical and membrane processes^{11,12,23-25} and facilitates synaptic plasticity. Specifically, acetylcholine increased the magnitude and probability of induction of the long-term potentiation produced by high-frequency tetanic stimulation²⁶⁻²⁸. This facilitation of plasticity occurred under conditions that did not elicit rhythmic activity. Our results show that when rhythmic activity is induced, the changes in plasticity are even more dramatic. Under these conditions, high-frequency stimulation is no longer required and long-term enhancement is reliably produced even by weak stimulation that would normally have no effect. Thus an important role of cholinergic modulation may be to put the hippocampal networks into a highly plastic oscillatory state that is specialized for learning. This may help to explain why the learning of new information is so strongly disrupted by anticholinergic drugs, such as scopolamine, or by the loss of cholinergic neurons, as in Alzheimer's dementia. □

Overexpression of dystrophin in transgenic *mdx* mice eliminates dystrophic symptoms without toxicity

Gregory A. Cox*, Nell M. Cole†, Kiihiro Matsumura‡, Stephanie F. Phelps*, Stephen D. Hauschka§, Kevin P. Campbell†, John A. Faulkner† & Jeffrey S. Chamberlain*||

* Department of Human Genetics, † Institute of Gerontology, || Human Genome Center, University of Michigan Medical School, Ann Arbor, Michigan 48109-0618, USA

‡ Howard Hughes Medical Institute and Department of Physiology and Biophysics, University of Iowa College of Medicine, Iowa City, Iowa 52242, USA

§ Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

DUCHENNE and Becker muscular dystrophy (DMD and BMD) are X-linked recessive diseases caused by defective expression of dystrophin^{1,2}. The *mdx* mouse, an animal model for DMD, has a mutation that eliminates expression of the 427K muscle and brain isoforms of dystrophin^{1,3,4}. Although these animals do not display overt muscle weakness or impaired movement, the diaphragm muscle of the *mdx* mouse is severely affected and shows progressive myofibre degeneration and fibrosis which closely resembles the human disease^{5,6}. Here we explore the feasibility of gene therapy for DMD by examining the potential of a full-length dystrophin transgene to correct dystrophic symptoms in *mdx* mice. We find that expression of dystrophin in muscles of transgenic *mdx* mice eliminates the morphological and immunohistological symptoms of muscular dystrophy. In addition, overexpression of dystrophin prevents the development of the abnormal mechanical properties associated with dystrophic muscle without causing deleterious side effects. Our results provide functional evidence for the feasibility of gene therapy for DMD.

The dystrophin complementary DNA expression vector (pMDA) used to generate transgenic *mdx* mice is diagrammed in Fig. 1a. Tissue-specific expression of this full-length murine dystrophin cDNA⁷ was controlled by regulatory regions of the mouse muscle creatine kinase (MCK) gene^{8,9}. Immunoblot analysis of transgenic *mdx* muscle tissue revealed that dystrophin expression from this transgene was ~50 times the level of endogenous dystrophin in control C57BL/10 muscles (Fig. 1b). The exogenous protein in transgenic *mdx* muscle comigrated with control dystrophin and migrated at a position corresponding to $M_r \sim 400K$. Dystrophin transgene expression was specific to skeletal and cardiac muscle tissues and was not detected in uterine smooth muscle or in non-muscle tissues such as liver and brain (Fig. 1b). A smear of dystrophin immunostaining is seen in the undiluted transgenic skeletal and cardiac muscle lanes and is presumably due to detection of excess dystrophin breakdown products. The overexpressed transgenic dystrophin, normally 0.002% of total muscle protein¹, was also observed as a prominent 400K band in Coomassie-blue-stained SDS-polyacrylamide gels, but was not visible in control or *mdx* muscle extracts (Fig. 1b).

A variety of assays demonstrated that the high levels of dystrophin produced in the transgenic mice were sufficient to eliminate dystrophic symptoms from *mdx* muscles. The localization of dystrophin by immunofluorescence in control, *mdx* and transgenic *mdx* muscle is shown in Fig. 2a. Increased dystrophin expression was apparent from the increased immunofluorescence of transgenic *mdx* muscles compared with control muscles. Dys-

Received 9 March; accepted 18 May 1993.

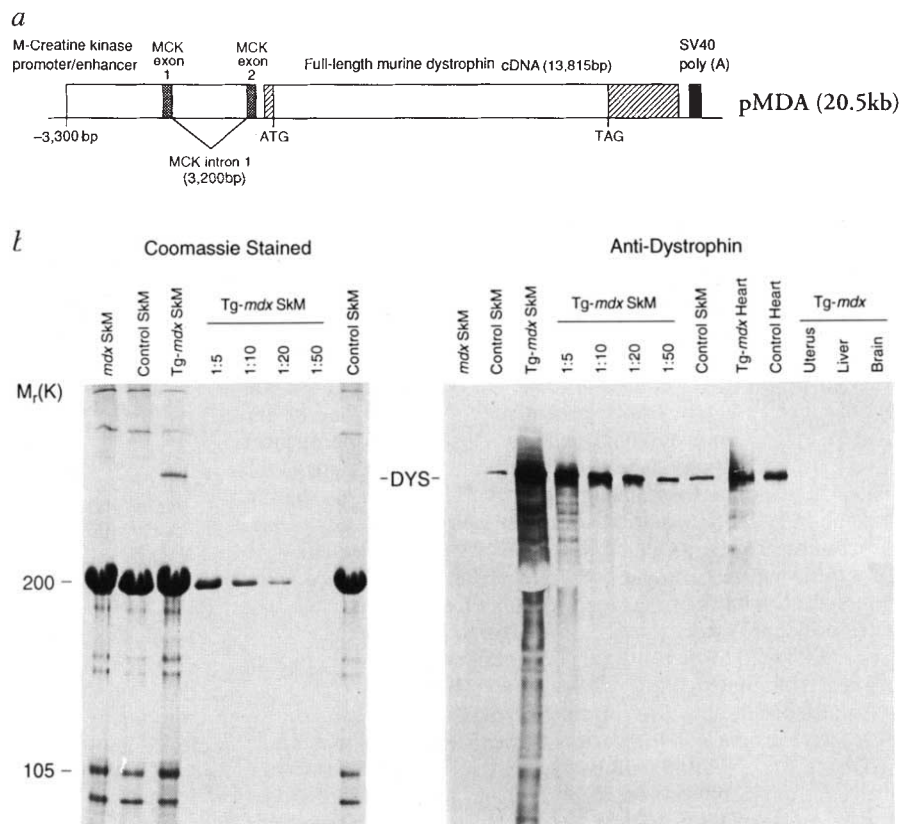
1. Fibiger, H. C. *Trends Neurosci.* **14**, 220-223 (1991).
2. Aigner, T. G. & Mishkin, M. *Behav. neural Biol.* **45**, 81-87 (1986).
3. Coyle, A. E., Price, D. L. & DeLong, M. R. *Science* **219**, 1184-1190 (1983).
4. Bland, B. H. *Prog. Neurobiol.* **26**, 1-54 (1986).
5. Squire, L. R. *Memory and Brain* (Oxford Univ. Press, 1987).
6. Winson, J. *Science* **201**, 160-163 (1978).
7. Otto, T., Eichenbaum, H., Wiener, S. I. & Wible, S. G. *Hippocampus* **1**, 181-192 (1991).
8. Konopacki, J., MacIver, M. B., Bland, B. H. & Roth, S. H. *Brain Res.* **405**, 196-198 (1987).
9. Konopacki, J., Bland, B. H. & Roth, S. H. *Brain Res.* **455**, 110-114 (1988).
10. MacVicar, B. A. & Tse, F. W. Y. *J. Physiol., Lond.* **417**, 197-212 (1989).
11. Dutar, P. & Nicoll, R. A. *J. Neurosci.* **8**, 4214-4224 (1988).
12. Nicoll, R. A., Malenka, R. C. & Kauer, J. A. *Physiol. Rev.* **70**, 513-565 (1990).
13. Colom, L. V., Nassif-Caudarella, S., Dickson, C. T., Smythe, J. W. & Bland, B. H. *Hippocampus* **1**, 381-390 (1991).
14. Eichenbaum, H., Kuperstein, M., Fagan, A. & Nagode, J. J. *Neurosci.* **7**, 716-732 (1987).
15. Traub, R. D., Miles, R. & Buzsáki, G. *J. Physiol., Lond.* **451**, 653-672 (1992).
16. Linás, R. R. *Science* **242**, 1654-1664 (1988).
17. Gray, C. M., König, P., Engel, A. K. & Singer, W. *Nature* **338**, 334-337 (1989).
18. Pavlidis, C., Greenstein, Y. J., Grudman, M. & Winson, J. *Brain Res.* **439**, 383-387 (1988).
19. Larson, J., Wong, D. & Lynch, G. *Brain Res.* **368**, 347-350 (1986).
20. Stanton, P. K. & Sejnowski, T. J. *Nature* **339**, 215-218 (1989).
21. Rose, G. M. & Dunwiddie, T. V. *Neurosci. Lett.* **69**, 244-248 (1986).
22. Christie, B. R. & Abraham, W. C. *Neuron* **9**, 79-84 (1992).
23. Halliwell, J. V. *Prog. Brain Res.* **84**, 255-272 (1990).
24. Markram, H. & Segal, M. J. *Physiol., Lond.* **427**, 381-393 (1990).
25. Markram, H. & Segal, M. J. *Physiol., Lond.* **447**, 513-533 (1992).
26. Burgard, E. C. & Sarvey, J. M. *Neurosci. Lett.* **116**, 34-39 (1990).
27. Schulz, P. E. & Johnston, D. *Soc. Neurosci. Abstr.* **16**, 655 (1990).
28. Brocher, S., Artola, A. & Singer, W. *Brain Res.* **573**, 27-36 (1992).
29. Alger, B. E. et al. in *Brain Slices* (ed. Dingle, R.) 381-437 (Plenum, New York, 1984).

ACKNOWLEDGEMENTS. This work was supported by a grant from the NIH.

|| To whom correspondence should be addressed.

FIG. 1 a, Full-length dystrophin cDNA expression vector pMDA. The expression of the murine dystrophin cDNA⁷ was regulated by 6.5 kb of MCK gene sequences, including -3,300 bp upstream of exon 1, the complete first intron, and a truncated exon 2 deleted just 5' of the MCK initiator methionine⁸. A muscle specific enhancer is located 1,200 bp upstream of the transcription start site and a modulatory control element is located within the first intron⁹. The SV40 polyadenylation site was inserted 3' of the dystrophin cDNA. Non-coding regions of dystrophin cDNA are represented by hatched boxes. b, Coomassie-blue-stained 6% SDS-polyacrylamide gel and immunoblot analysis of dystrophin (DYS) expression in SDS-extracts of 6-month C57BL/10 normal control, *mdx* and transgenic *mdx* (Tg-*mdx*) mouse muscle. Molecular weight standards ($\times 10^{-3}$) are indicated on the left. SkM: skeletal muscle.

METHODS. Transgenic mice were generated by microinjection of purified pMDA insert into F₂ hybrid zygotes from C57BL/6J $\times$ SJL/J parents as described²². Positive transgenic mice were identified by the polymerase chain reaction (PCR) using primers specific for MCK intron 1 and the dystrophin cDNA. Two founder F₀ females were generated and upon mating with C57BL/10 *mdx* males, one of the F₀ mice transmitted the transgene to F₁ offspring. Transgenic mice on the *mdx* background were produced by breeding transgenic F₁ females heterozygous for the *mdx* mutation with C57BL/10 *mdx* males. The progeny were screened for the presence of the transgene by PCR and for the presence of the *mdx* point mutation by an allele-specific oligonucleotide assay using PCR primers flanking murine exon 23 as described²³. Dystrophin expression was analysed by loading 100 μ g of total protein from each tissue and 1/5, 1/10, 1/20 and 1/50 dilutions of transgenic *mdx* skeletal muscle extracts (as determined by the Coomassie-plus protein assay (Pierce)) onto identical



6%-polyacrylamide gels. One gel was stained with Coomassie brilliant blue and the other transferred to nitrocellulose and stained with Dys-2 monoclonal antibody (Nova Castra), which recognizes the final 17 amino acids of dystrophin. Electrophoretic transfer to a membrane and immunostaining have been described⁴.

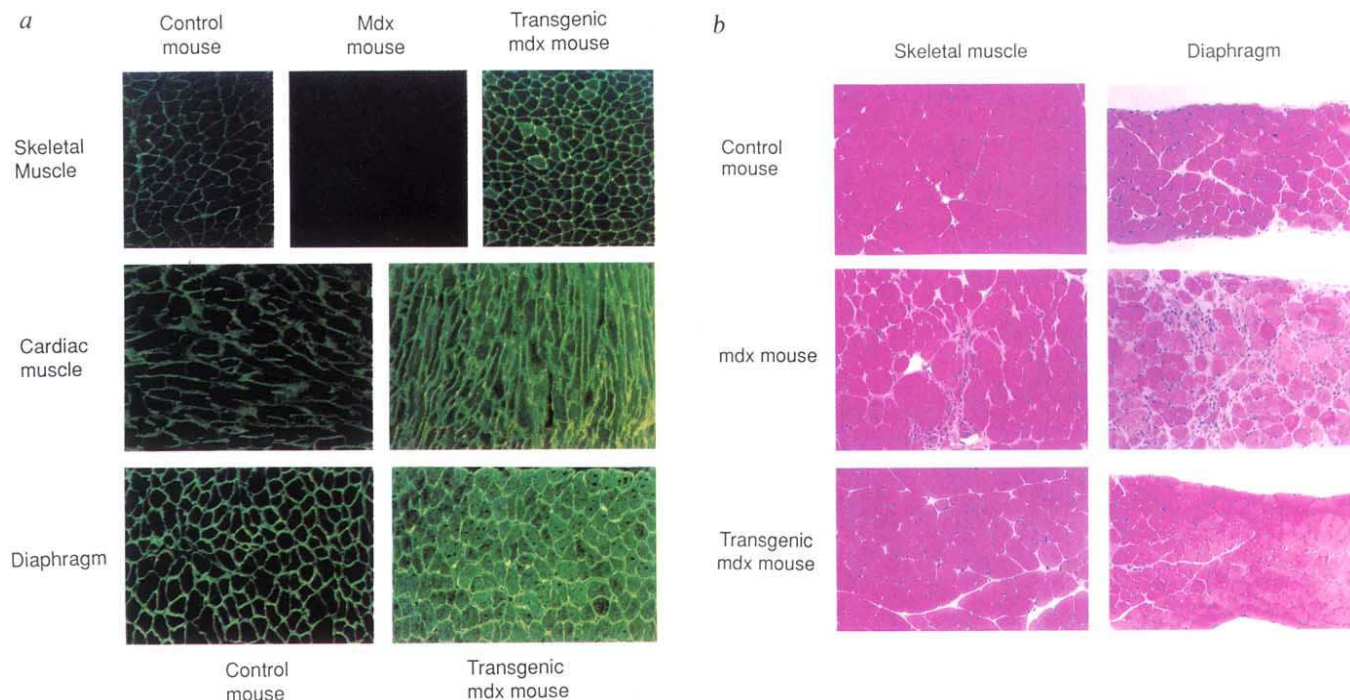


FIG. 2 Immunostaining and histological analysis of age-matched (6-month) C57BL/10 control, *mdx* and transgenic *mdx* muscle. a, Dystrophin immunostaining is localized to the sarcolemma of control quadriceps, cardiac and diaphragm muscles and is absent from *mdx* muscle, except for rare revertant dystrophin-positive fibres. (Magnifications: ▶

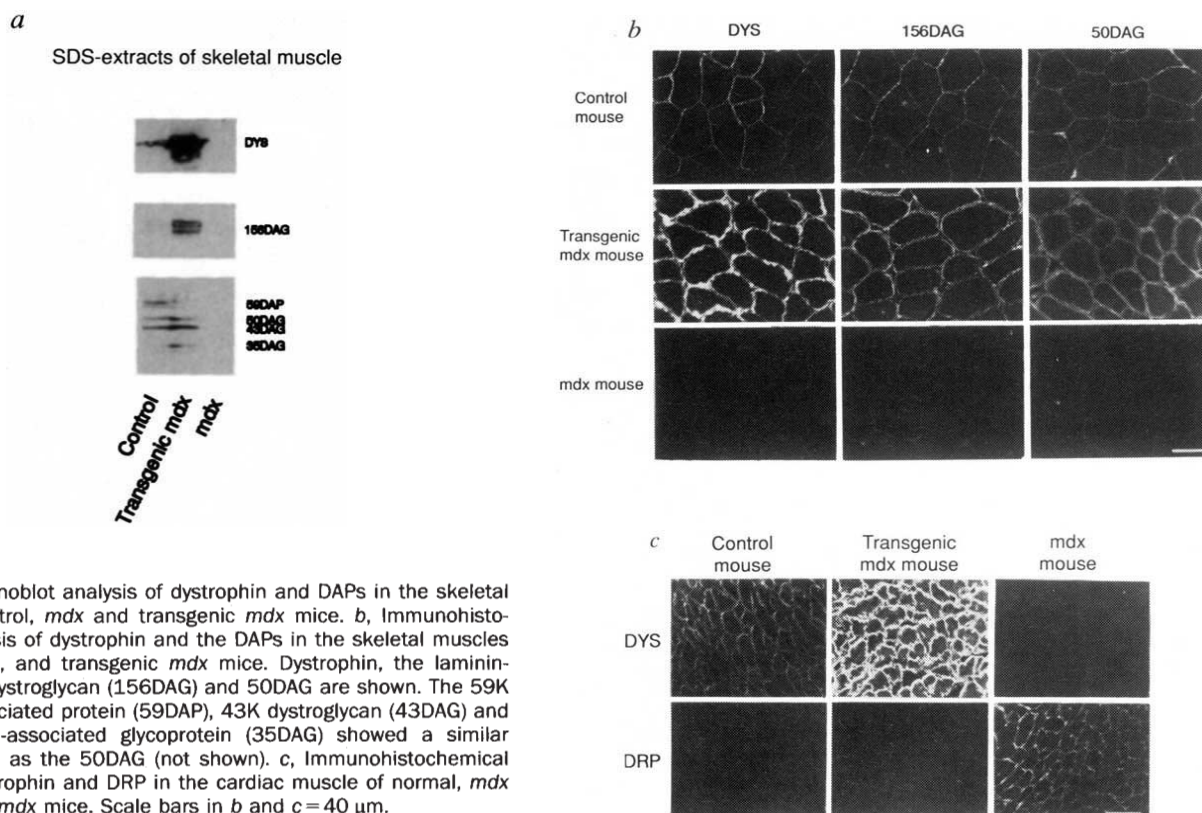


FIG. 3 *a*, Immunoblot analysis of dystrophin and DAPs in the skeletal muscles of control, *mdx* and transgenic *mdx* mice. *b*, Immunohistochemical analysis of dystrophin and the DAPs in the skeletal muscles of normal, *mdx*, and transgenic *mdx* mice. Dystrophin, the laminin-binding 156K dystroglycan (156DAG) and 50DAG are shown. The 59K dystrophin-associated protein (59DAP), 43K dystroglycan (43DAG) and 35K dystrophin-associated glycoprotein (35DAG) showed a similar staining pattern as the 50DAG (not shown). *c*, Immunohistochemical analysis of dystrophin and DRP in the cardiac muscle of normal, *mdx* and transgenic *mdx* mice. Scale bars in *b* and *c* = 40 μ m.

METHODS. Immunoblot analysis of SDS extracts of muscle specimens^{15,16} and immunohistochemistry of muscle specimens¹³⁻¹⁶ have been described. Affinity-purified rabbit antibody against the last 10 amino acids of the C terminus of dystrophin (*a*), affinity-purified rabbit antibody against the first 15 amino acids of the N terminus of dystrophin

(*b* and *c*), affinity-purified rabbit antibody against the last 12 amino acids of the C terminus of DRP (*c*), monoclonal antibody against 156DAG (*a*) and affinity-purified sheep antibodies against 156DAG, 59DAP, 50DAG, 43DAG and 35DAG (*a*, *b*) were used^{13-16,24,25}.

trophin expressed from the transgene was localized correctly to the sarcolemma, although additional dystrophin staining in the sarcoplasm of transgenic *mdx* muscles was observed and may have been due to saturation of membrane-binding sites. Histological examination of control, *mdx* and transgenic *mdx* limb skeletal muscles and diaphragm muscles revealed a complete correction of dystrophic pathology in the transgenic animals (Fig. 2*b*). Features of *mdx* muscle pathology such as necrosis, large variation in fibre size, increased degeneration and regeneration with centrally located nuclei, and progressive degeneration and fibrosis of diaphragm muscle, were all absent from the transgenic *mdx* mice.

quadriceps, 11.6 $\times$; cardiac and diaphragm, 23 $\times$). *b*, Haematoxylin- and eosin-stained sections of transgenic *mdx* quadriceps and diaphragm muscles show corrected dystrophic *mdx* histopathology. Note the absence of fibrosis and centrally located nuclei in the control and transgenic *mdx* muscles (magnifications, 23 $\times$). The apparent difference in fibre size between transgenic and control muscle was not reproduced in further sections and is not a general feature of transgenic *mdx* muscle.

METHODS. Immunostaining of unfixed 7 μ m muscle cryosections was done with a 1/1,000 dilution of a rabbit polyclonal antibody against the first 410 amino acids of dystrophin using conditions described previously⁴. Histological 4 μ m sections were prepared from muscle tissues fixed in 4% paraformaldehyde and 1% glutaraldehyde, embedded in glycol methacrylate and stained with haematoxylin and eosin. Samples were photographed with brightfield illumination on a Nikon Optiphot-2 microscope equipped with a UFX-DX photomicrographic system.

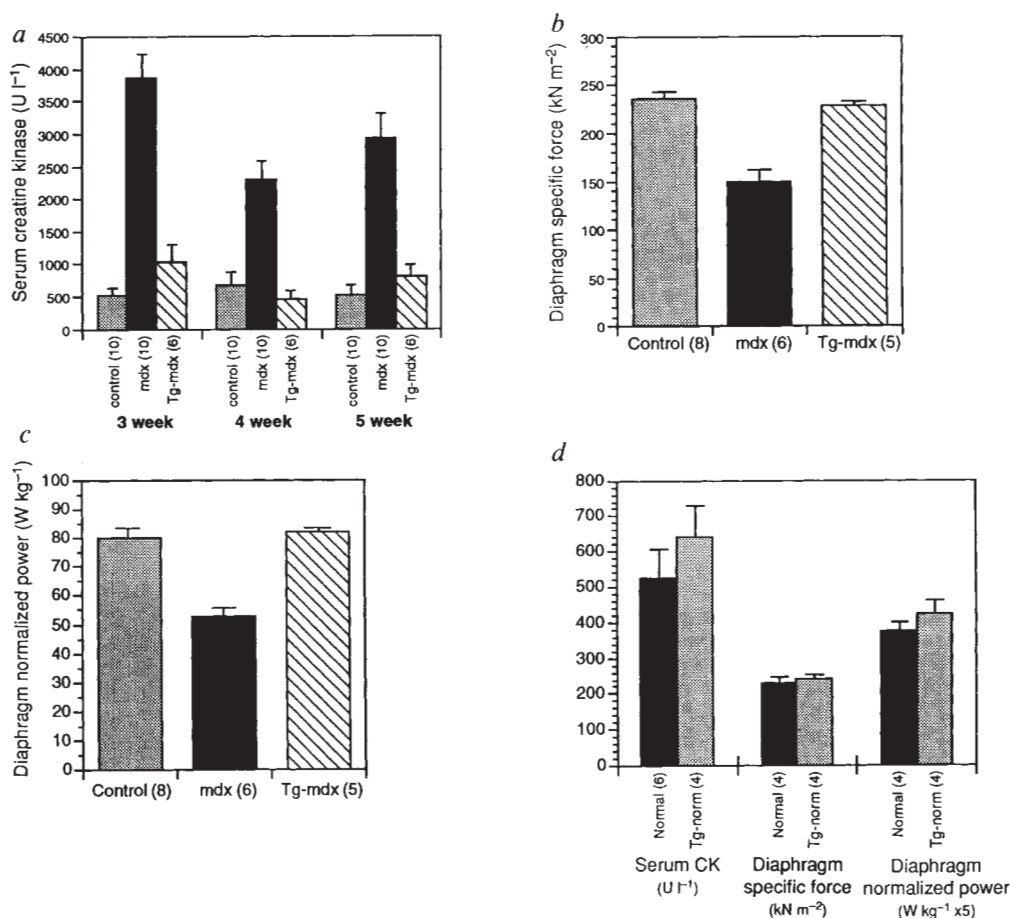
Dystrophin is normally associated with a large oligomeric complex of sarcolemmal glycoproteins, including dystroglycan, which provides a linkage to the extracellular matrix component laminin^{10,12}. In the absence of dystrophin in DMD patients and *mdx* mice, all of the dystrophin-associated proteins (DAPs) are reduced in the sarcolemma^{13,14}. The specific deficiency of the 50K dystrophin-associated glycoprotein (50DAG) alone causes a muscular dystrophy with a DMD-like phenotype¹⁵. These findings suggest that all of the DAPs, in addition to dystrophin, must be restored in the sarcolemma for the transfer of the dystrophin gene to be successful in correcting the *mdx* phenotype. Immunoblot analysis of skeletal muscle extracts demonstrates the restoration of all of the DAPs in transgenic *mdx* skeletal muscles (Fig. 3*a*). Immunohistochemical analysis of limb (Fig. 3*b*) and diaphragm (data not shown) skeletal muscles reveals that both dystrophin and the DAPs co-localize to the sarcolemma in the transgenic *mdx* mouse. Interestingly, both immunoblot and immunohistochemical analysis (Figs 3*a*, *b*) indicate that the ratio of dystrophin to DAP expression is greater in transgenic muscle than in control muscle. Furthermore, immunoblot analysis of skeletal muscle membrane preparations reveals that the amount of dystrophin in the membranes appear to be the same for transgenic *mdx* and control mice (data not shown). These findings suggested that the sarcolemma of transgenic *mdx* mice was saturated with the DAPs and that a substantial fraction of the overexpressed dystrophin was not associated with the DAPs. The sarcolemmal expression of the dystrophin-related protein (DRP) has been reported to be upregulated in *mdx* cardiac muscle¹⁶. We were unable to detect DRP by immunohistochemical analysis of the sarcolemma of cardiac

FIG. 4 Dystrophin expression in skeletal and cardiac muscles of transgenic *mdx* mice results in control levels of serum creatine kinase (CK). Serum samples were assayed at 3, 4 and 5 weeks of age in F_3 normal heterozygous (+/*mdx*) female, and hemizygous *mdx* and transgenic *mdx* (Tg-*mdx*) male littermates (26 animals in 3 F_3 litters). The serum CK levels of heterozygous (+/*mdx*) mice are not significantly different from wild type, enabling these littermates to be used as normal controls¹⁷. The number of mice (*n*) in each group is shown in parentheses. *b* and *c*, Specific forces and normalized powers developed by diaphragm muscles from a sample of the same littermates as in *a* at 3 months of age. *d*, Comparison of serum CK levels, diaphragm specific force and diaphragm normalized power in phenotypically normal F_3 heterozygous (+/*mdx*) female littermates with (Tg-norm) or without (normal) the dystrophin transgene. Serum CK values were based on an average value of each animal measured at 3, 4 and 5 weeks of age.

METHODS. Serum samples from mice were obtained from the retro-orbital sinus using heparinized capillary tubes. Serum CK measurements were determined using a coupled enzyme spectrophotometric assay (Sigma). Diaphragm strips 1 to 2 mm wide, including an adjacent section of a single rib and part of the central tendon, were cut from the central region of the lateral costal hemidiaphragm and immersed in an oxygenated bath containing mammalian Ringer solution (pH 7.4) at 25 °C. Muscles were adjusted to the optimum lengths (L_0) for the development of isometric force. Force was determined during maximum isometric tetanic contractions. Power output was determined by isovelocity shortenings from 100% L_0 to 90% L_0 during maximum muscle activation. Initiation of the isovelocity shortening ramp and stimulation of the muscle occurred simultaneously.

muscles from transgenic *mdx* and control mice (Fig. 3c). In contrast, DRP was detected in the sarcolemma of cardiac muscles from the *mdx* mouse (Fig. 3c). The suppression of the upregulation of DRP in cardiac muscles from transgenic *mdx* mice was confirmed by immunoblot analysis (data not shown).

The physiological effects of dystrophin expression on muscle pathology were assessed by measuring serum creatine kinase levels at 3, 4 and 5 weeks of age in control, *mdx*, and transgenic *mdx* littermates (Fig. 4a). Heterozygous (+/*mdx*) female mice display an extremely limited pathology^{17,18} and are thus useful as age-matched phenotypically normal controls. There were no significant differences in the serum creatine kinase levels of control and transgenic *mdx* littermates (*t*-test with unequal variances, $P > 0.05$), but serum creatine kinase was significantly higher in *mdx* mice than in the transgenic *mdx* littermates ($P < 0.001$) at all times tested. Functional correction of the *mdx* muscle pathology was demonstrated by comparing the contractile performance of *in vitro* bundles of diaphragm muscle fibres from control, *mdx* and transgenic *mdx* littermates at 3 months of age (Figs 4b and c). Compared with the mean values for control mice of 236 kN m^{-2} for specific force and 79.9 W kg^{-1} for normalized power, the values for muscles from transgenic *mdx* mice were not significantly different, whereas those for muscles from *mdx* mice were significantly lower at 63 and 66% respectively ($P < 0.001$). In addition, there were no differences in



Stimulation was terminated at the end of the shortening ramp. Power output during a single contraction was calculated as the product of average force and velocity of shortening. The velocity of shortening and the frequency of stimulation were adjusted to elicit maximum power output^{26,27}. After measurements of power, the central tendon and rib bone were trimmed and the muscle was blotted and weighed immediately. Specific force (kN m^{-2}) values were normalized to mean cross-sectional area. Power (W) was normalized by muscle mass (W kg^{-1}).

serum creatine kinase, and normalized force and power of the diaphragm muscle between normal (+/*mdx*) female littermates with and without the dystrophin transgene ($P > 0.05$; Fig. 4d), and so no toxic side effects resulted from overexpression of dystrophin.

We have demonstrated the functional correction of muscular dystrophy by tissue-specific expression of a full-length dystrophin cDNA clone in transgenic *mdx* mice. Although the high levels of dystrophin obtained may partially reflect the site of transgene integration, the MCK promoter is known to be highly active in skeletal and cardiac muscle⁹. Furthermore, MCK-regulated transgenes expressing other dystrophin isoforms have also given high levels of muscle expression (J.S.C. *et al.*, unpublished). These studies indicate that MCK regulatory regions may be useful in the design of vectors for muscle gene therapy. A transgenic *mdx* mouse expressing extremely low amounts of a mutant dystrophin from a murine retroviral promoter has been reported¹⁹. These mice had their symptoms partially corrected, but it was unclear if the absence of a complete correction was due to low protein expression or to functional deficiencies of the truncated dystrophin clone. The correct size of the transgenic protein on immunoblots and detection with antibodies against the N terminus and C terminus (Figs 1b and 3b), indicate that our transgene encodes the entire dystrophin protein. Furthermore, overexpression of the full-length dystrophin transgene does not

result in an adverse phenotype (Figs 2b and 4d). This lack of toxicity suggests that tight control over dystrophin expression levels may not be necessary for correction of muscular dystrophy. Delivery of exogenous dystrophin by viral vectors with sub-optimal infection frequencies might benefit from overexpression of dystrophin in portions of a multinucleated myofibre without concern for cytotoxicity²⁰. Although progressive skeletal muscle weakness and wheelchair confinement are the most obvious symptoms of DMD, most patient deaths result from respiratory failure²¹. The restoration of normal values for force and power development in transgenic *mdx* diaphragm muscles provides functional evidence supporting the feasibility of gene therapy for DMD. □

Received 23 April; accepted 18 June 1993.

- Hoffman, E. P., Brown, R. H. Jr & Kunkel, L. M. *Cell* **51**, 919–928 (1987).
- Ahn, A. H. & Kunkel, L. M. *Nature Genet.* **3**, 283–291 (1993).
- Sicinski, P. et al. *Science* **244**, 1578–1580 (1989).
- Cox, G. A., Phelps, S. F., Chapman, V. M. & Chamberlain, J. S. *Nature Genet.* **4**, 87–93 (1993).
- Partridge, T. *Neuropathol. appl. Neurobiol.* **17**, 353–363 (1991).
- Stedman, H. H. et al. *Nature* **352**, 536–539 (1991).
- Lee, C. C., Peariman, J. A., Chamberlain, J. S. & Caskey, C. T. *Nature* **349**, 334–336 (1991).
- Jaynes, J. B., Chamberlain, J. S., Buskin, J. N., Johnson, J. E. & Hauschka, S. D. *Molec. cell Biol.* **6**, 2855–2864 (1986).
- Johnson, J. E., Wold, B. J. & Hauschka, S. D. *Molec. cell Biol.* **9**, 3393–3399 (1989).
- Ervasti, J. M., Ohlndieck, K., Kahl, S. D., Gaver, M. G. & Campbell, K. P. *Nature* **345**, 315–319 (1990).

- Ervasti, J. M. & Campbell, K. P. *Cell* **66**, 1121–1131 (1991).
- Ibraghimov-Beskrovnaya, O. et al. *Nature* **355**, 696–702 (1992).
- Ohlndieck, K. & Campbell, K. P. *J. Cell Biol.* **115**, 1685–1694 (1991).
- Ohlndieck, K. et al. *Neurology* **43**, 795–800 (1993).
- Matsumura, K. et al. *Nature* **359**, 320–322 (1992).
- Matsumura, K., Ervasti, J. M., Ohlndieck, K., Kahl, S. & Campbell, K. P. *Nature* **360**, 588–591 (1992).
- Chapman, V. M., Miller, D. M., Armstrong, D. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1292–1296 (1989).
- Watkins, S. C., Hoffman, E. P., Slayter, H. S. & Kunkel, L. M. *Muscle Nerve* **12**, 861–868 (1989).
- Wells, D. J. et al. *Hum. molec. Genet.* **1**, 35–40 (1992).
- Ragot, T. et al. *Nature* **361**, 647–650 (1993).
- Emery, A. E. H. *Duchenne Muscular Dystrophy* (Oxford Monog. Med. Genet. No. 15, Oxford Medical, 1987).
- Hogan, B., Constantini, F. & Lacey, E. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1986).
- Chamberlain, J. S., Phelps, S. F., Cox, G. A., Maichele, A. J. & Greenwood, A. D. in *Molecular and Cell Biology of Muscular Dystrophy* (ed. Partridge, T.) 167–189 (Chapman & Hall, London, 1993).
- Ohlndieck, K. et al. *Neuron* **7**, 499–508 (1991).
- Ervasti, J. M., Kahl, S. D. & Campbell, K. P. *J. biol. Chem.* **266**, 9161–9165 (1991).
- Faulkner, J. A., Zerbe, E. & Brooks, S. V. *Am. J. Physiol.* **259**, 259–265 (1990).
- Brooks, S. V. & Faulkner, J. A. *J. Physiol.* **436**, 701–710 (1991).

ACKNOWLEDGEMENTS. We thank K. Corrado, J. Rafael and J. Buskin for advice and discussions, and P. Mills, J. Wang, R. Bräbäck, A. Christensen and the Transgenic Animal Model Core, University of Michigan, for technical assistance. This work was supported by grants from the Muscular Dystrophy Association (U.S.A.) (to J.S.C. and S.D.H.); by grants from The National Institute of Arthritis and Musculoskeletal Diseases (to J.S.C. and S.D.H.); The National Institute of Diabetes, Digestive and Kidney Diseases (to J.S.C.); The National Heart, Lung and Blood Institute (to S.D.H.); and The National Institute on Aging (to J.A.F.). NIH, G.A.C. is supported by an NSF graduate fellowship; K.M. is supported by a post doctoral fellowship from the Muscular Dystrophy Association (U.S.A.); J.S.C. is a recipient of a Basil O'Connor Starter Scholar Research Award from The March of Dimes Birth Defects Foundation; K.P.C. is an Investigator of The Howard Hughes Medical Institute.

Helper T-cell development in the absence of CD4-p56^{lck} association

Nigel Killeen & Dan R. Littman

Howard Hughes Medical Institute, and Departments of Microbiology and Immunology and of Biochemistry and Biophysics, University of California, San Francisco, California 94143–0414, USA

THE CD4 and CD8 glycoproteins are expressed on helper and cytotoxic T lymphocytes, respectively, and have important functions in the differentiation and activation of these cells^{1–8}. These molecules are thought to participate in signal transduction by binding to the same class II or class I major histocompatibility complex molecules that are engaged by the T-cell antigen receptor^{9–13}. The cytoplasmic domains of both CD4 and CD8 interact with the protein tyrosine kinase p56^{lck} (refs 14–17), an essential participant in thymocyte maturation¹⁸ and T-cell activation¹⁹. This interaction is required for effective *in vitro* responses to antigen^{5–8}, suggesting that signalling through p56^{lck} is a major function of CD4 and CD8. Here we investigate the role of the CD4-p56^{lck} interaction during T-lymphocyte development by expressing wild-type and truncated products of CD4 transgenes in mice that lack endogenous CD4 and hence have defective helper-cell development^{3,4}. We find that transgenic CD4, which cannot associate with p56^{lck}, can nevertheless rescue the helper-cell lineage when overexpressed. This result indicates that the contribution of CD4 to lineage development need not involve signalling through p56^{lck}, and provides insight into the general function of CD4 and CD8.

Transgenic mice expressing wild-type or cytoplasmic domain-deleted CD4 molecules were generated using murine CD4 minigenes^{20,21} under the transcriptional regulation of the CD3δ enhancer and promoter²² (Fig. 1). These animals were then bred for homozygosity of a null mutation at the endogenous CD4 locus⁴, thereby permitting analysis of the function of the transgene-encoded proteins. To assess the extent of lineage development in the various transgenic mice, thymocytes and lymph node

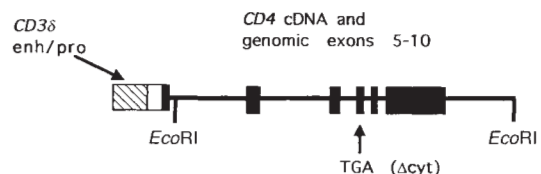


FIG. 1 Structure of the CD4 transgenes. Minigene constructs^{20,21} encoding wild-type or cytoplasmic domain-deleted (Δ cyt) CD4 molecules were fused to the enhancer (enh) and promoter (pro) elements of the mouse CD3δ gene (NotI-Sall fragment of pNeZ; ref. 22). The same minigenes under transcriptional regulation of the *lck* proximal promoter have been described previously^{20,21}. Promoter substitution was effected in order to facilitate an evaluation of the influence of transgene expression on phenotype. CD3δ-CD4 transgenes were microinjected³⁴ into C57BL/6 × SJL F₁ eggs that had been fertilized by CD4^{-/-} males (C57BL/6 × DBA/2 × 129). Founders and their progeny were identified by FACS analysis of peripheral blood mononuclear cells using anti-CD4 (GK1.5-PE from Becton-Dickinson). One line expressing wild-type CD4 and seven expressing tailless CD4 were then back crossed to CD4^{-/-} mice.

cells were analysed by flow cytometry (Fig. 2). As previously shown⁴, in the absence of CD4 expression (Fig. 2a, $-/-$), the development of the CD4 lineage was hampered and only ~10% of peripheral T cells were CD8⁺, as compared with 60–70% CD4⁺CD8⁺ T cells in mice expressing endogenous CD4 (Fig. 2a, $+/-$). Expression of the wild-type transgene, at a level similar to that of endogenous CD4, re-established the normal proportions of CD4 and CD8 lineage cells, which could be readily discriminated by their differential expression of CD8 (Fig. 2a, $-/-$ wt). Introduction of CD4 lacking the cytoplasmic tail into the CD4^{-/-} background also resulted in rescue of the CD8⁺ T-cell subset, but the number of these cells varied between different transgenic lines. Strikingly, there was a direct correlation between the proportion of CD4 lineage cells and the level of cell-surface expression of the tailless CD4 molecule. Analyses of three transgenic lines expressing different levels of the tailless molecule are shown in Fig. 2a–c ($-/-\Delta$ cyt). Whereas expression of the tailless molecule at a level lower than that of endogenous

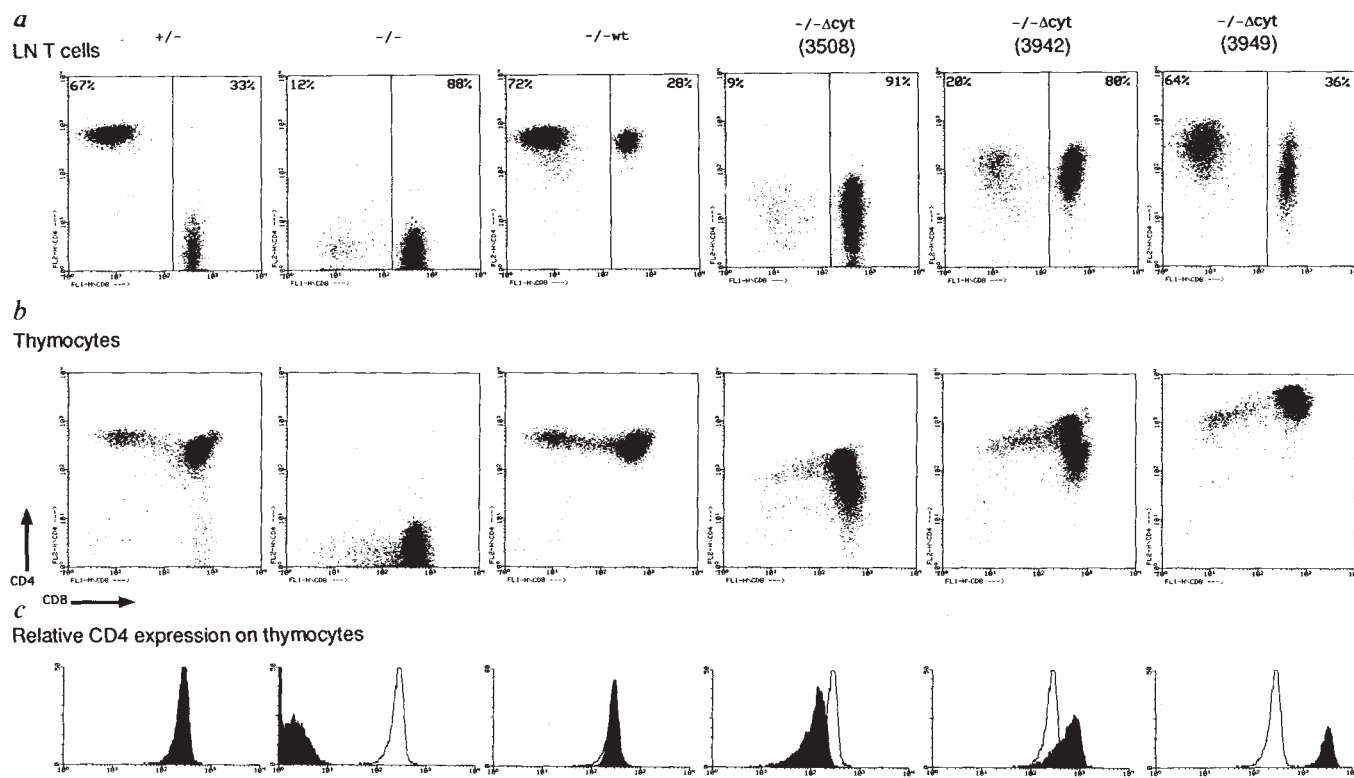


FIG. 2 Expression of transgenic CD4 in CD4-deficient mice. Dot plots show expression of CD4 and CD8 on lymph node (LN) cells (a) and thymocytes (b) from mice of the indicated genotypes. c, Filled histograms show expression of CD4 on immature (CD3^{duil}) thymocytes relative to the level on cells from the control CD4^{+/-} mouse. d, Graph of the relationship between mean relative fluorescence of tailless CD4 on CD3^{duil} thymocytes and the percentage of CD4-lineage peripheral T cells in individual CD4Δcyt transgenic lines. Similar phenotypes were also observed for CD4^{-/-} mice bearing *lck*-CD4 transgenes (not shown). METHODS. Thymocytes and lymph node cells were stained with anti-CD4, anti-CD8 (GK1.5-PE and 53-6.7-FITC from Becton-Dickinson) and anti-CD3ε(145-2C11-biotin from Pharmingen) followed by streptavidin PE-Cy5 (Caltag Tricolor). Cells were analysed using a Becton-Dickinson FACSscan and Lys II software. Dot plots for peripheral lymphocytes and histograms for CD3^{duil} thymocytes were generated by gating for expression of CD3.

CD4 did not alter the ratios of CD4⁺CD8⁻ to CD4⁺CD8⁺ cells, five- to sixfold overexpression restored the development of the CD4 lineage to normal levels (Fig. 2, number 3949). In total, five different transgenic founders expressing the tailless CD4 molecule were analysed in the CD4-negative background; this pattern was notably consistent (Fig. 2d).

Previous studies have shown that a cysteine-containing motif within the cytoplasmic domain of CD4 is required for binding to p56^{lck} (refs 15, 16). To confirm that deletion of this domain of CD4 in the reconstituted animals eliminated the CD4-p56^{lck} association, anti-CD4 immunoprecipitates from thymocyte lysates were tested for associated p56^{lck} by *in vitro* kinase assay and by immunoblot with a rabbit antiserum specific for a p56^{lck} amino-terminal peptide (Fig. 3). Immunoprecipitates from cells expressing endogenous or transgenic wild-type CD4 included a protein of 55–60 K that was recognized by the *lck*-specific antiserum (Fig. 3a). These immunoprecipitates also contained abundant kinase activity both for an exogenous substrate, enol-

ase, and for a protein running at the expected size of p56^{lck} (Fig. 3b). Immunoprecipitates from CD4-deficient thymocytes and from those overexpressing the tailless variant of CD4 had negligible kinase activity and no detectable p56^{lck} by immunoblot (Fig. 3a and b). Tailless CD4 was abundant in these immunoprecipitates, as shown by parallel immunoblot with a rabbit antiserum specific for mouse CD4 (not shown).

To determine whether the appearance of CD4 lineage cells resulted in the expected restoration of helper T-cell function, transgenic mice were immunized with a T-cell dependent antigen, TNP-KLH, and tested for their ability to generate a primary TNP-specific antibody response. As previously reported, the mean TNP-specific antibody titre in the sera of mice deficient in CD4 expression was ~10-fold lower than that in mice expressing endogenous CD4 (ref. 4). This defect was largely corrected by expression of either the wild-type or tailless transgene-encoded CD4 molecules (Fig. 4a). In mixed lymphocyte cultures, T cells from CD4-deficient mice are poor responders to stimulation with

cells from major histocompatibility complex (MHC) congenic mice expressing the MHC class II alloantigen I-A^{bm12}, but they respond well to cells bearing the class I alloantigen K^{bm1} (refs 3, 4; Fig. 4b). Mice reconstituted with wild-type as well as tailless CD4 were corrected in their response to bm12 (Fig. 4b). As expression of transgenic CD4 in CD8 lineage cells renders them responsive to the bm12 alloantigen²³, both CD4 and CD8 lineage cells can proliferate in this assay. Although our results are consistent with rescue of helper T lymphocytes by tailless CD4, it remains to be determined whether the tailless molecule is as effective as its wild-type counterpart in the rescued peripheral T cells.

Results from several different experimental systems have suggested that CD4 and CD8 have signal transducing properties that are important for T-cell lineage determination²⁴⁻²⁶, selection^{20,21} and activation⁵⁻⁸. Signalling has been proposed to involve the associated p56^{lck} because the interaction of CD4 and CD8 with this protein tyrosine kinase (PTK) is required for effective activation of antigen-specific T-cell hybridomas⁵⁻⁸. We show here that in T-cell development CD4 can function independently of association with p56^{lck}, albeit less efficiently than the wild-type molecule. Therefore, the CD4-associated PTK does not appear to have an essential signalling role in the selection of the helper-T-cell lineage.

The direct correlation between cell-surface concentration of tailless CD4 and the extent of lineage rescue suggests that, rather than providing a direct signalling function, the primary role of CD4 is in adhesion, stabilizing the interaction of the T-cell antigen receptor (TCR) with antigen-MHC class II complexes. At limiting concentrations of CD4, its interaction with p56^{lck} may be critical for recruitment into the TCR complex, where simultaneous engagement of MHC class II would add stability to an otherwise weak interaction^{27,28}. This requirement for association with the PTK can be circumvented when tailless CD4 is overexpressed, presumably because this improves the probability that CD4 will encounter the same MHC class II molecule that is engaged by the TCR. This model of co-receptor function is sup-

ported by several findings. For example, the CD4 cytoplasmic domain is required for its colocalization with the TCR complex following TCR stimulation^{7,29,30}. Furthermore, TCR stimulation affects the avidity of CD8 for immobilized MHC class I (ref.

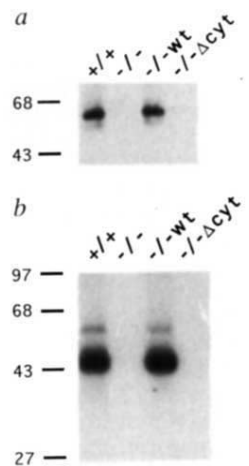


FIG. 3 Absence of p56^{lck} association with transgene-encoded CD4 lacking the cytoplasmic domain. Immunoblot (a) and *in vitro* kinase assay (b) of anti-CD4 immunoprecipitates from lysates of thymocytes. M, calibration is indicated on the left ($\times 10^{-3}$). METHODS. 5×10^7 thymocytes from mice of the indicated genotypes ($-/-$ -wt and $-/-$ - Δ cyt 3949, as in Fig. 2) were incubated on ice for 30 min with anti-CD4 (GK1.5), pelleted and then lysed in 1 ml LB (15 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM PMSF, 1 mM vanadate, 1 μ g ml⁻¹ aprotinin, 1% Brij96). Lysates were clarified by centrifugation and then incubated with goat anti-rat 1gG-agarose (Sigma) for 30 min. Beads were pelleted and washed sequentially with LB/0.5 M LiCl, LB and finally kinase assay buffer (20 mM Tris-HCl, pH 7.4, 10 mM MnCl₂, 0.1% Brij96). Half of each immunoprecipitate was taken for immunoblot with a rabbit antiserum specific for an amino-terminal peptide of p56^{lck} (a). The other half of each immunoprecipitate was tested for kinase activity (b) as follows³⁵. Agarose beads were resuspended in 20 μ l kinase assay buffer containing 1.5 μ l γ -[³²P]ATP ($>5,000$ Ci mmol⁻¹) and 4 μ g enolase and incubated at room temperature for 10 min. Reactions were stopped by adding 20 μ l of 2 $\times$ non-reducing sample buffer before heat denaturation and SDS-PAGE. After electrophoresis, gels were fixed, treated with 1 M KOH for 1 h at 55 $^{\circ}$ C and then refixed before drying and autoradiography. Similar results were also obtained for immunoprecipitates of NP-40 lysates of thymocytes and also for cells from the equivalent *lck*-CD4 transgenic lines (not shown).

ported by several findings. For example, the CD4 cytoplasmic domain is required for its colocalization with the TCR complex following TCR stimulation^{7,29,30}. Furthermore, TCR stimulation affects the avidity of CD8 for immobilized MHC class I (ref.

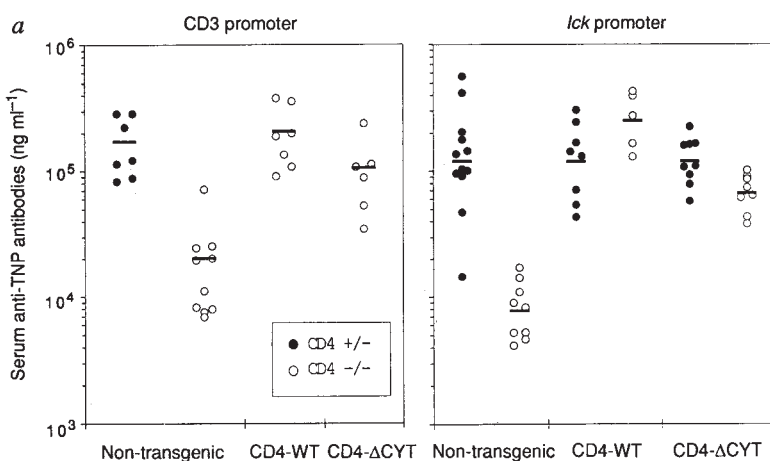
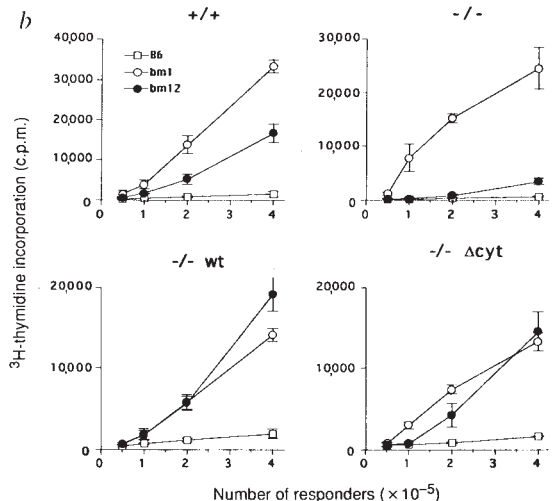


FIG. 4 Helper cell function in transgene-reconstituted CD4-deficient mice. a, Restoration of antibody response. Mice were immunized intraperitoneally with 10 μ g trinitrophenyl-keyhole limpet haemocyanin (TNP-KLH) in alum and sera were drawn on days 0, 10 and 14. Plots show serum TNP-specific antibody titres on day 10 measured by indirect ELISA as previously described⁴. The mice used for this experiment were transgenic for CD4 minigenes under the control of either the *lck* proximal or CD3 δ promoters. Mice bearing the wild-type form of the *lck*-CD4 transgene had roughly normal proportions of CD4 and CD8 lineage peripheral T cells, whereas the *lck*-promoter-CD4 Δ cyt transgenic mice had slightly less than half of the normal proportion of CD8⁺ T cells (not



shown). CD3 promoter lines are the same as those shown in Fig. 2 ($-/-$ -wt and $-/-$ - Δ cyt 3949). b, Restoration of allo-class II reactivity in mixed lymphocyte cultures. Lymph node cells from H-2^b mice of the indicated genotypes were challenged *in vitro* with irradiated anti-Thy-1-depleted spleen cells from C57BL/6 mice or from mice of the two MHC congenic strains B6.C-H-2^{bm1}/ByJ (MHC class I mutation) and B6.C-H-2^{bm12}/KhEg (MHC class II mutation). ³H-thymidine was added after 3 d and the assay was collected 18 h later. CD4 $-/-$ -wt and CD4 $-/-$ - Δ cyt lines are the same as those shown in Fig. 2 ($-/-$ -wt and $-/-$ - Δ cyt 3949). Similar results were also obtained when the equivalent *lck*-CD4 transgenic lines were tested for bm12 reactivity (not shown).

31), which would presumably improve the ability of CD8 to stabilize TCR-MHC class I interactions. A similar effect for CD4-MHC class II adhesion may also occur. Finally, the hypothesis that p56^{lck} promotes CD4-TCR colocalization is supported by results indicating that the CD4-associated p56^{lck} molecule need not have kinase activity to potentiate the function of CD4 (ref. 36).

The rescue of CD4-lineage helper cells by the tailless molecule is most consistent with a stochastic mechanism for the development of CD4⁺CD8⁻ cells from CD4⁺CD8⁺ precursor thymocytes^{32,33}. If the decision to silence CD8 gene expression and embark on the CD4-helper developmental pathway is made in a random fashion, then a signal mediated through the CD4 cytoplasmic tail would not be required and the stabilization function outlined above would suffice to drive positive selection of class II-restricted cells. It remains possible, however, that the mutant CD4 protein can still transmit an instructive signal, perhaps through association of its transmembrane domain with other signalling components²⁶.

Unselected thymocytes that will eventually commit to the CD4 lineage may have variable dependency on the function of CD4, possibly determined by the relative affinity of their TCRs for antigen/MHC class II. We have recently shown that, in the absence of CD4, some class II-restricted helper T cells do complete development and have a CD8⁺TCR $\alpha\beta$ ⁺ phenotype (ref. 37). The developmental rescue by tailless CD4 described here may perhaps be confined to such CD4-independent cells or to cells that have a weak requirement for the co-receptor function. Preliminary analysis of TCR variable domain usage by CD8⁺TCR $\alpha\beta$ ⁺ T cells in CD4^{-/-} mice shows normal diversity with no particular bias, and there are no obvious phenotypic characteristics that distinguish these cells. Uncovering the extent to which individual CD4⁺ T cells depend on the function of CD4 will probably necessitate a better understanding of the role of receptor-ligand affinity in shaping the T-cell repertoire. □

35. Beyers, A. D., Spruyt, L. L. & Williams, A. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2945-2949 (1992).

36. Xu, H. & Littman, D. R. *Cell* (in the press).

37. Locksley, R., Reiner, S., Hatam, F., Littman, D. R. & Killeen, N. *Science* (in the press).

ACKNOWLEDGEMENTS. We thank K. Forbush, A. Garvin and R. Perimutter for *lck*-CD4 transgenic mice; R. Coffman and members of his laboratory for help with ELISAs; C. Crooks, S. Sawada and S. Stuart for constructs and for assistance in the generation of the CD3 δ -CD4 transgenic mice; K. Chu, D. Law, H. Xu and M. Hill for advice and assistance with the biochemical assays; P. Johnson for rabbit antiserum against mouse p56^{lck}, and R. Locksley, M. Bix and C. Davis for comments on the manuscript. This work was supported by the Howard Hughes Medical Institute.

Coated vesicle assembly in the Golgi requires only coatamer and ARF proteins from the cytosol

Lelio Orci*, David J. Palmer†, Mylène Amherdt* & James E. Rothman†

* Department of Morphology, University of Geneva Medical Center, 1 rue Michel-Servet, 1211 Geneva, Switzerland

† Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

TRANSPORT vesicles derived from the Golgi apparatus are thought to mediate biosynthetic transport across the Golgi stack¹⁻³. These vesicles are surrounded by a protein coat⁴ whose principal constituents are coatamer (a complex of seven distinct subunits or COPs)⁵ and ADP-ribosylation factor (ARF, an *N*-myristylated small GTP-binding protein). The coat proteins of the COP-coated vesicles were originally defined by ultrastructural criteria⁴⁻⁹, however, and it is possible that important but minor coat proteins or cytoplasmic proteins needed for coat assembly may have been overlooked. Here we show that coatamer and ARF are the only cytoplasmic proteins needed for the assembly and budding of COP-coated vesicles. COP-coated buds may therefore form essentially by self-assembly from Golgi cisternae after an initial step in which GTP is used to allow ARF binding.

Figure 1a is an electron micrograph of a membrane pellet from incubations in which coatamer was bound to Golgi membranes in an ARF- and GTP- γ S-dependent reaction (in the presence of ATP). Numerous coated buds and vesicles are present, and analysis of serial sections (an example is shown in Fig. 1) reveals that 60-70% of the coated profiles are fully formed coated vesicles, the remainder being coated buds. Coated profiles are missing when either coatamer (Fig. 1b) or ARF proteins, or both (not shown) are omitted from the binding reaction.

Immunoelectron microscopy (Fig. 2) shows that coatamer localized with antibody against the β -COP subunit (Fig. 2a) and ARF (Fig. 2b) are present in the coated buds and vesicles. In the reaction shown in Fig. 1a, 85 $\pm$ 2% of the β -COP associated with Golgi membranes is present on coated buds and vesicles, the remainder being found on cisternal membranes outside morphologically recognizable coats. The coated buds and vesicles also contain 47 $\pm$ 3% of the membrane-bound ARF, the remainder being found on bulk cisternal membranes. The bulk of the material produced by this simple and well defined coatamer binding reaction is thus made up of coated buds and coated vesicles. Coatamer and ARF are therefore not only necessary but also sufficient to form coated vesicles from Golgi membranes, and no major coat proteins required to form vesicles can have escaped notice in earlier compositional studies^{4,5}.

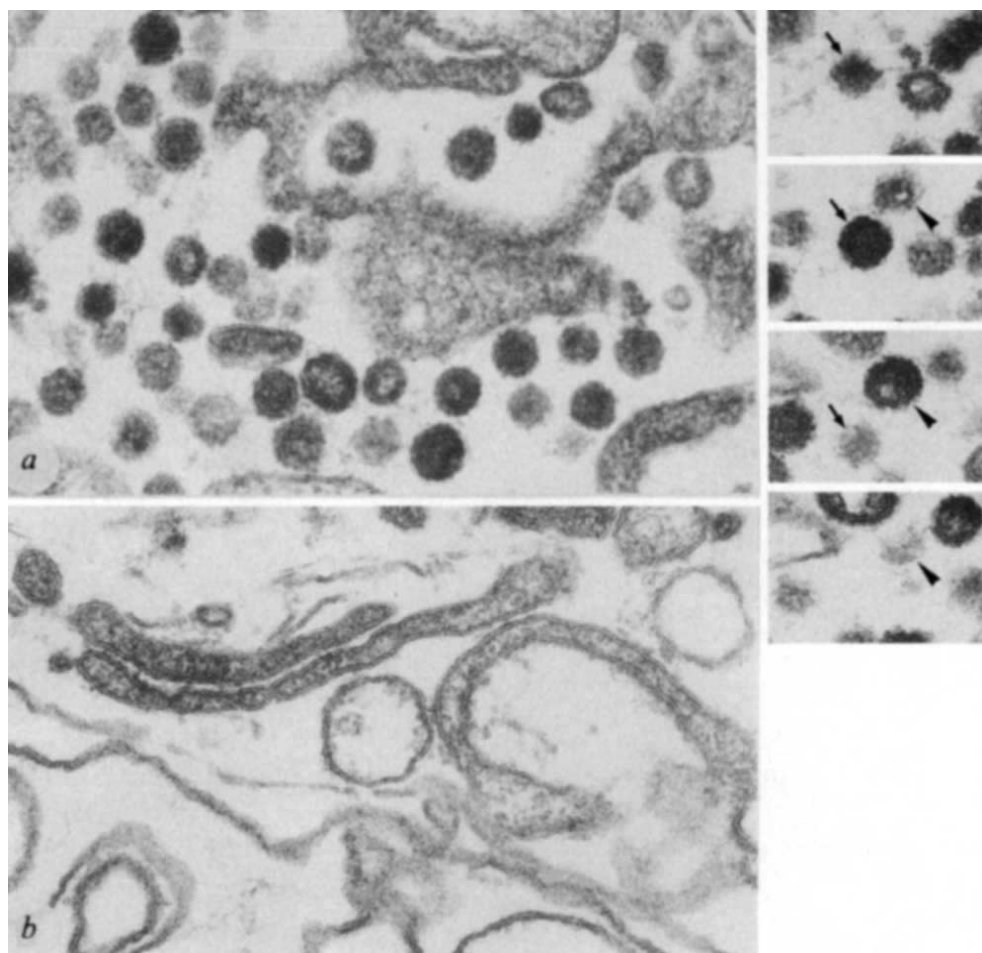
Although not necessary, do cytoplasmic components other than ARF and coatamer affect the coat assembly and budding reaction? To test this, we used a cytosol fraction that was depleted of ~85% of its endogenous coatamer by means of a

Received 15 March; accepted 10 June 1993.

1. Fowlkes, B. J. & Pardoll, D. M. *Adv. Immun.* **44**, 207-264 (1989).
2. Fung-Leung, W. P. et al. *Cell* **65**, 443-449 (1991).
3. Rahemtulla, A. et al. *Nature* **353**, 180-183 (1991).
4. Killeen, N., Sawada, S. & Littman, D. R. *EMBO J.* **12**, 1547-1553 (1993).
5. Zomoyiska, R. et al. *Nature* **342**, 278-281 (1989).
6. Glaichenhaus, N., Shastri, N., Littman, D. R. & Turner, J. M. *Cell* **64**, 511-520 (1991).
7. Collins, T. L. et al. *J. Immun.* **148**, 2159-2162 (1992).
8. Micelli, M. C., von Hoegen, P. & Parnes, J. R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2623-2627 (1991).
9. Janeway, C. A. Jr. *Rev. Immun.* **10**, 645-674 (1992).
10. Salter, R. D. et al. *Nature* **345**, 41-46 (1990).
11. Aldrich, C. J. et al. *Nature* **352**, 718-721 (1991).
12. Ingold, A. L., Landel, C., Knall, C., Evans, G. A. & Potter, T. A. *Nature* **352**, 721-723 (1991).
13. Killeen, N., Moriarty, A., Lacy, H. S. & Littman, D. R. *J. exp. Med.* **175**, 89-97 (1992).
14. Rudd, C. E., Trevillyan, J. M., Dasgupta, J. D., Wong, L. L. & Schlossman, S. F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5190-5194 (1988).
15. Shaw, A. S. et al. *Molec. cell. Biol.* **10**, 1853-1862 (1990).
16. Turner, J. M. et al. *Cell* **60**, 755-765 (1990).
17. Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. *Cell* **55**, 301-308 (1988).
18. Molina, T. J. et al. *Nature* **357**, 161-164 (1992).
19. Straus, D. B. & Weiss, A. *Cell* **70**, 585-593 (1992).
20. Teh, H. S. et al. *Nature* **349**, 241-243 (1991).
21. van Oers, N. S. C. et al. *Eur. J. Immun.* **22**, 735-743 (1992).
22. Lee, N. A., Loh, D. Y. & Lacy, E. J. *exp. Med.* **175**, 1013-1025 (1992).
23. Robey, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 608-612 (1991).
24. Robey, E. et al. *Cell* **64**, 99-107 (1991).
25. Borgulya, P., Kishi, H., Muller, U., Kirberg, J. & von Boehmer, H. *EMBO J.* **10**, 913-918 (1991).
26. Seong, R. H., Chamberlain, J. W. & Parnes, J. R. *Nature* **356**, 718-720 (1992).
27. Kiyoshi, M. et al. *Science* **254**, 1788-1791 (1993).
28. Weber, S., Trautwein, A., Oliveri, F., Gerhard, W. & Karjalainen, K. *Nature* **356**, 793-796 (1992).
29. Rojo, J. M., Saizawa, K. & Janeway, C. A. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3311-3315 (1989).
30. Mittler, R. S., Goldman, S. J., Spitalny, G. L. & Burakoff, S. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8531-8535 (1989).
31. O'Rourke, A. M., Rogers, J. & Mescher, M. F. *Nature* **346**, 187-189 (1990).
32. Chan, S. H., Cosgrove, D., Waltzinger, C., Benoist, C. & Mathis, D. *Cell* **73**, 225-236 (1993).
33. Davis, C. B., Killeen, N., Crooks, M. E. C., Raulet, D. & Littman, D. R. *Cell* **73**, 237-247 (1993).
34. Hogan, B. L. M., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory, New York, 1986).

FIG. 1 a, Field of a Golgi preparation incubated in the presence of ARF, GTP- γ S and coatomer. Abundant coated vesicular profiles are present. The inserts on the right show four consecutive serial sections that demonstrate the existence of free coated vesicles. Given the vesicle's diameter (75–80 nm) and the thickness of the sections ($\sim$ 70–80 nm), a free vesicle is captured in one equatorial or near-equatorial view (that is, with the bilayer of the limiting membrane clearly distinguishable) in one section; the next two sections show grazing views on each side of the equator. None of the equatorial or grazing views shows a connection with a cisternal element, hence the 'free' nature of the vesicle. One such vesicle on consecutive sections is identified by the small arrows; another vesicle is labelled with arrowheads. b, Field of a Golgi preparation incubated as in a, but without coatomer. No accumulation of coated profiles is evident. See Table 1 for quantitation. Magnification: in a, 78,900 $\times$; in b, 82,200 $\times$.

METHODS. Rabbit liver Golgi membranes⁸ (not salt-washed) (190 μ g protein per ml) were incubated in 200 μ l in 0.2 M sucrose, 25 mM KCl, 20 mM HEPES-KOH, pH 7.2, 2.5 mM magnesium acetate, 20 μ M GTP- γ S and 1 mg ml⁻¹ soybean trypsin inhibitor in the presence or absence of coatomer (isolated from bovine liver, using gel filtration as the final purification step^{5,22}; 10 μ g ml⁻¹), myristoylated ARF-1 (purified from an *E. coli* strain expressing both human ARF-1 and yeast *N*-myristyl transferase^{23,24}; 40 μ g ml⁻¹), bovine brain cytosol⁸ immunodepleted of coatomer¹⁰ (1.8 mg ml⁻¹), 50 μ M ATP and ATP-regenerating system (2 mM creatine phosphate, 8 IU ml⁻¹ creatine kinase). Incubations were for 20 min at 37 °C. Membranes were collected by centrifugation, fixed with glutaraldehyde, contrasted with tannic acid and embedded in Epon. Thin sec-



tions were further contrasted with uranyl acetate and lead citrate and photographed in a Philips CM10 electron microscope. Photographs taken at calibrated magnification were printed on paper. The quantitation of coated vesicles and coated buds was carried out with the aid of an electronic pen and appropriate computer program, in a double-blind fashion using a code that was not broken until quantitation was complete.

monoclonal anti-coatomer IgG (ref. 10) (following coatomer removal by western blotting with antibody to the β -COP subunit) as a source of endogenous ARF proteins, as well as any other unknown factors that might affect coated vesicle assembly and budding. Immunodepleted cytosol can replace recombinant ARF-1 protein in the budding reaction (Table 1), but does not significantly affect the yield of either coated vesicles or coated buds when added in addition to ARF-1 protein. The effects of coatomer depletion upon reconstituted intercisternal transport are entirely consistent with the proposed role of these vesicles as carriers¹¹ and will be described elsewhere.

Assembly of the COP coat and bud formation occur in a process resembling self-assembly¹², in that only the coat proteins themselves are required from the cytoplasm. Although clathrin and adaptors can self-assemble in solution into coats, less is known about the mechanism of *de novo* formation of clathrin-coated vesicles from plasma or Golgi membranes¹³. Clathrin will self-assemble into coated pits, however, when added back to stripped plasma membranes^{14,16} and functional endocytosis can be reconstituted by adding back partially purified clathrin and purified adaptors to permeabilized cells¹⁷. Binding of clathrin adaptors to the Golgi, like coatomer, requires an ARF protein in the GTP-bound form¹⁸, indicating that clathrin adaptor binding is probably a prerequisite for clathrin coat assembly. An

obvious difference between COP and clathrin coats is that the coat and adaptor functions are separate in the latter, but form part of the same complex (coatomer) in the former. The significance of this difference is unclear, but it is likely that the coatomer fulfils the structural roles played by both the clathrin

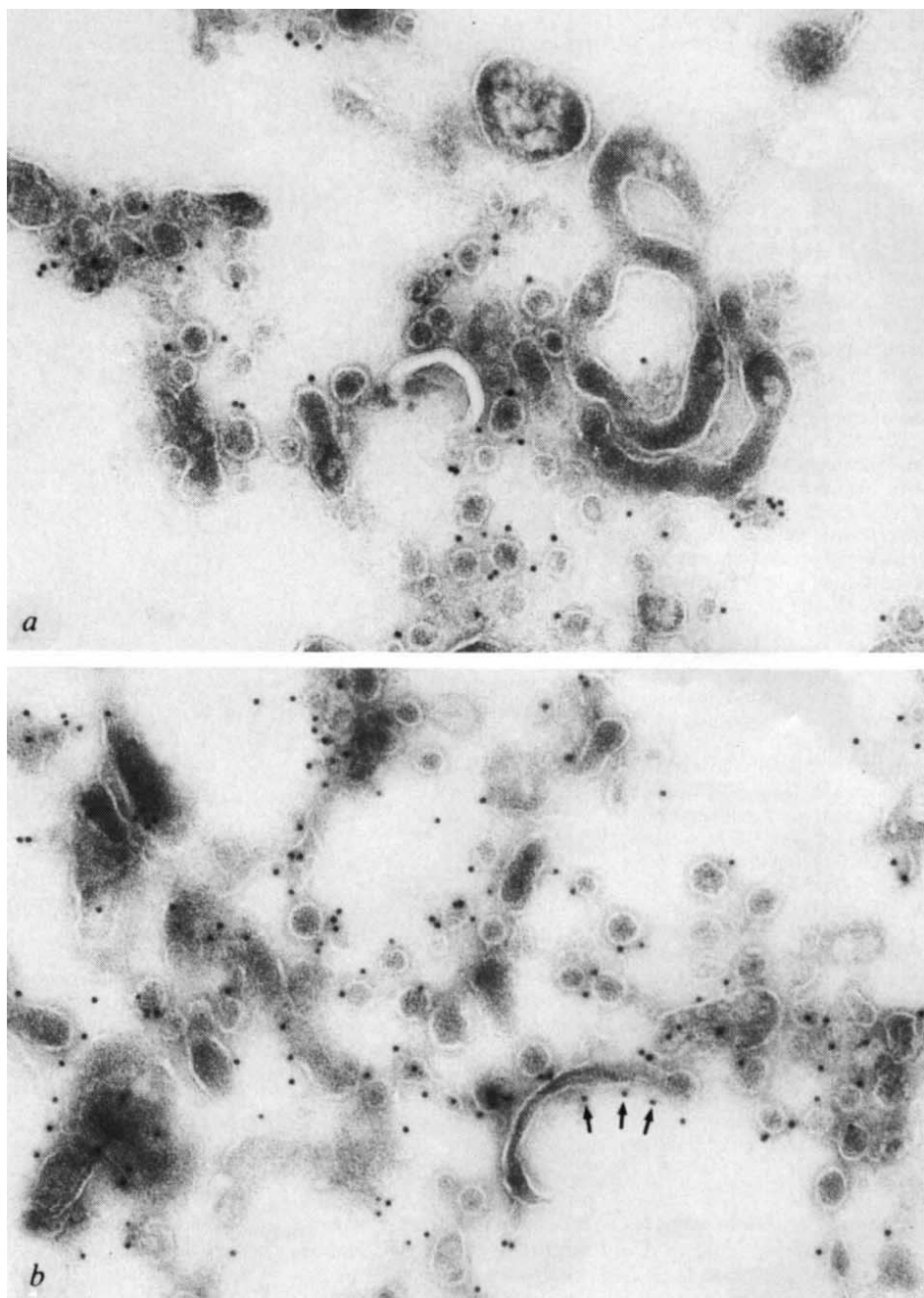
TABLE 1 Requirements for coated vesicle assembly

Coatomer	ARF	Cytosol depleted of coatomer	Density of coated buds plus coated vesicles (number per μ m ²)
+	—	—	2.2 $\pm$ 0.5
—	+	—	1.6 $\pm$ 0.4
—	—	+	3.3 $\pm$ 0.6
+	+	—	23.8 $\pm$ 1.4
+	—	+	32.1 $\pm$ 1.4
+	+	+	22.0 $\pm$ 1.3
—	+	+	1.7 $\pm$ 0.5

Rabbit liver Golgi membranes were incubated in the presence (indicated by a plus sign) or absence of coatomer, recombinant ARF-1 protein, or cytosol depleted of coatomer. All incubations contained GTP- γ S and ATP and ATP-regenerating system. Electron microscopy and quantitation of coated vesicles and buds is described in Fig. 1 legend. A total of 30 Golgi areas were analysed from three different portions of each pellet.

FIG. 2 a, Thin cryosection of a Golgi preparation incubated with ARF, GTP- γ S and coatomer. The immunogold particles revealing β -COP predominate in association with buds and vesicles. b, Thin cryosection of a Golgi preparation incubated as in a, but immunostained with anti-ARF antibody. Gold particles are associated with vesicles, buds and flat cisternal membrane (the latter indicated by arrows). Magnification in a, and b, 78,900 $\times$.

METHODS. Membranes, incubated as described in Fig. 1 legend, were pelleted, fixed with glutaraldehyde and cryosectioned as described²⁵. The protein A-gold method²⁶ was used to reveal antibodies on sections. Dilution of the anti- β -COP antibody (affinity-purified anti-EAGE; gift from T. Kreis) was 1:40. Dilution of the anti-ARF antibody (2048, raised against the whole ARF protein)²⁰ was 1:100. The protein A-gold solution was diluted 1:100. Quantitation of immunogold labelling was carried out with the aid of an electronic pen and appropriate computer program, in a double-blind fashion using a code that was not broken until quantitation was complete.



and adaptor particles in clathrin-coated vesicles.

ARF has previously been shown to be necessary for coatomer to bind to Golgi membranes^{19,20}, but the resulting complexes were not shown to be assembled coats or their precursors. COP-coated vesicle assembly is inhibited by the N-terminal 15 residues of ARF-1 protein²¹, indicating that ARF might play a role in the process, but our results now provide direct evidence that it is required. □

Received 8 February; accepted 7 May 1993.

- Orci, L., Glick, B. S. & Rothman, J. E. *Cell* **46**, 171–184 (1986).
- Melancon, P. et al. *Cell* **51**, 1053–1062 (1987).
- Orci, L., Malhotra, V., Amherdt, M., Serafini, T. & Rothman, J. E. *Cell* **56**, 357–368 (1989).
- Serafini, T. et al. *Nature* **349**, 215–220 (1991).
- Waters, M. G., Serafini, T. & Rothman, J. E. *Nature* **349**, 248–251 (1991).
- Duden, R., Griffith, G., Frank, R., Argos, P. & Kreis, T. E. *Cell* **64**, 649–665 (1991).
- Serafini, T. et al. *Cell* **67**, 239–253 (1991).
- Malhotra, V., Serafini, T., Orci, L., Shepherd, J. C. & Rothman, J. E. *Cell* **58**, 329–336 (1989).
- Schekman, R. *Curr. Opin. Cell Biol.* **4**, 587–592 (1992).
- Orci, L. et al. *Nature* **362**, 648–652 (1993).

- Rothman, J. E. & Orci, L. *Nature* **355**, 409–415 (1992).
- Caspar, D. L. D. & Klug, A. *Cold Spring Harbor Symp. quant. Biol.* **27**, 1–24 (1962).
- Schmid, S. L. *BioEssays* **14**, 589–596 (1992).
- Moore, M. S., Mahaffey, D. T., Brodsky, F. M. & Anderson, R. G. W. *Science* **236**, 558–563 (1987).
- Mahaffey, D. T., Moore, M. S., Brodsky, F. M. & Anderson, R. G. W. *J. Cell Biol.* **108**, 1615–1624 (1989).
- Mahaffey, D. T., Peeler, J. S., Brodsky, F. M. & Anderson, R. G. W. *J. Biol. Chem.* **265**, 16514–16520 (1990).
- Smythe, E., Carter, L. L. & Schmid, S. L. *J. Cell Biol.* **119**, 1163–1171 (1992).
- Stammes, M. A. & Rothman, J. E. *Cell* **73**, 999–1005 (1993).
- Donaldson, J. G., Cassel, D., Kahn, R. A. & Klausner, R. D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6408–6412 (1992).
- Palmer, D. J., Helms, J. B., Beckers, C. J. M., Orci, L. & Rothman, J. E. *J. Biol. Chem.* **268**, 12083–12089 (1993).
- Kahn, R. A. et al. *J. Biol. Chem.* **267**, 13039–13046 (1992).
- Waters, M. G., Beckers, C. J. M. & Rothman, J. E. *Meth. Enzym.* **219**, 331–337 (1992).
- Weiss, O., Holden, J., Rulka, C. & Kahn, R. A. *J. Biol. Chem.* **264**, 21066–21072 (1989).
- Duronio, R. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1506–1510 (1990).
- Tokuyasu, K. T. *Histochem. J.* **12**, 381–403 (1980).
- Roth, J., Bendayan, M. & Orci, L. *J. Histochem. Cytochem.* **26**, 1074–1081 (1978).

ACKNOWLEDGEMENTS. This work was supported by a Swiss National Science Foundation grant to L.O. and an NIH grant to J.E.R., as well as a collaborative grant from the Human Frontier Science Program. D.J.P. was the recipient of an NIH Fogarty International Research Fellowship.

DNA topoisomerase V is a relative of eukaryotic topoisomerase I from a hyperthermophilic prokaryote

Alexei I. Slesarev*†, Karl O. Stetter‡, James A. Lake*, Martin Gellert§, Regis Krah§ & Sergei A. Kozyavkin†§

* Molecular Biology Institute and Department of Biology, University of California, Los Angeles, California 90024, USA

‡ Lehrstuhl für Mikrobiologie, Universität Regensburg, Universitätstrasse 31, 8400 Regensburg, Germany

§ Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Building 5, Bethesda, Maryland 20892-0002, USA

THE DNA topoisomerases are ubiquitous enzymes that fulfil vital roles in the replication, transcription and recombination of DNA by carrying out DNA-strand passage reactions¹⁻⁷. Here we characterize a prokaryotic counterpart to the eukaryotic topoisomerase I in the hyperthermophilic methanogen *Methanopyrus kandleri*⁸⁻¹⁰. The new enzyme, called topoisomerase V, has the following properties in common with eukaryotic topoisomerase I, which distinguish it from all other known prokaryotic topoisomerases: (1) its activity is Mg^{2+} -independent; (2) it relaxes both negatively and positively supercoiled DNA; (3) it makes a covalent complex with the 3' end of the broken DNA strand; and (4) it is recognized by antibody raised against human topoisomerase I. Eukaryotic-like enzymes have been discovered in some hyperthermophilic prokaryotes, namely the eocytes¹¹ and the extremely thermophilic archaeobacteria¹², and hyperthermophilic homologues of eukaryotic DNA polymerase- α , transcription factor IIB and DNA ligase¹³⁻¹⁵ have all been reported. Thus our findings support the idea that some essential parts of the eukaryotic transcription-translation and replication machineries were in place before the emergence of eukaryotes, and that the closest living relatives of eukaryotes may be hyperthermophiles.

Topoisomerase V was isolated and purified almost to homogeneity from *M. kandleri* using its ability to relax positively and negatively supercoiled DNA without magnesium (Fig. 1). According to the gel-filtration calibration standards, the peak of activity migrates as a globular protein with M_r 142,000, whereas SDS-PAGE gives an apparent M_r of 110,000 (Fig. 2b). This indicates that the topoisomerase is probably a monomer. From the yield of activity resulting from the purification procedure, we calculate that there are about 1,000 topoisomerase V molecules per *M. kandleri* cell.

The enzyme relaxes both positively and negatively supercoiled DNA at similar rates (Fig. 1). The final distribution of topoisomers does not depend on the initial supercoiling (Fig. 1a), and is not affected by a several-fold change in enzyme:DNA ratio (results not shown). In the conditions shown in Fig. 1a, the relaxation activity is catalytic and processive. The rate of DNA relaxation is proportional to the quantity of enzyme added and is about 15 supercoils per minute per enzyme monomer. DNA topoisomerase V does not require specific ions for activity: it relaxes DNA in potassium glutamate (up to 3.1 M) or in NaCl or KCl; EDTA, Mg^{2+} and ATP do not stimulate or inhibit the relaxation (results not shown).

The temperature dependence of topoisomerase V activity on positively and negatively supercoiled DNA is shown in Fig. 1b

and c. At 50 °C the enzyme has only weak activity. At 70–90 °C the enzymatic reaction yields fully relaxed duplex DNA. At temperatures above 90 °C, the enzyme produces highly unwound forms of DNA. This effect, named unlinking (that is, the substantial reduction in DNA linking number) is caused by DNA melting, and has been described in detail for *Desulfurococcus* topoisomerase III (ref. 16).

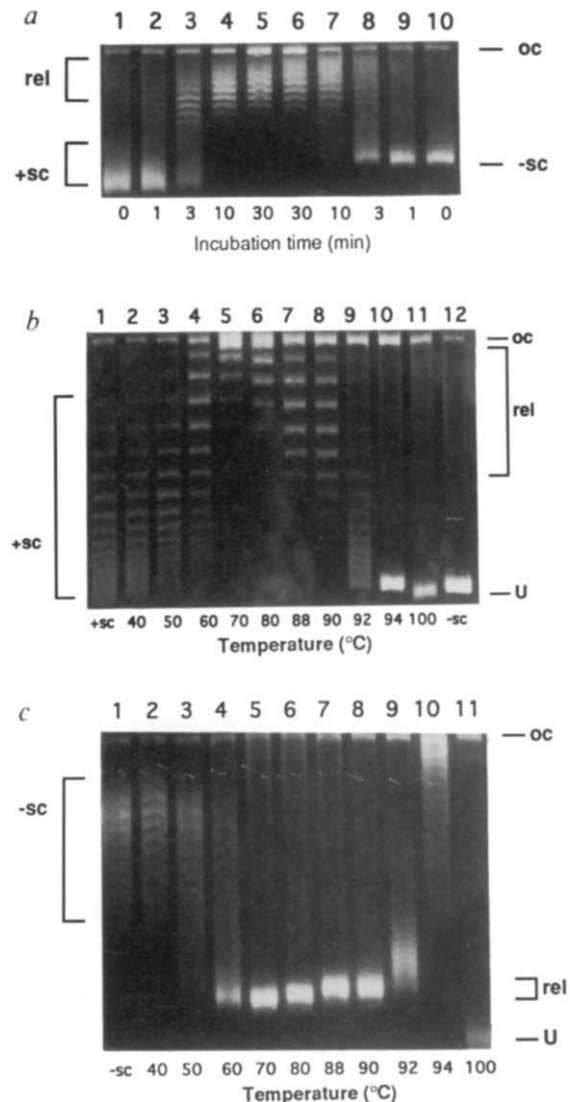


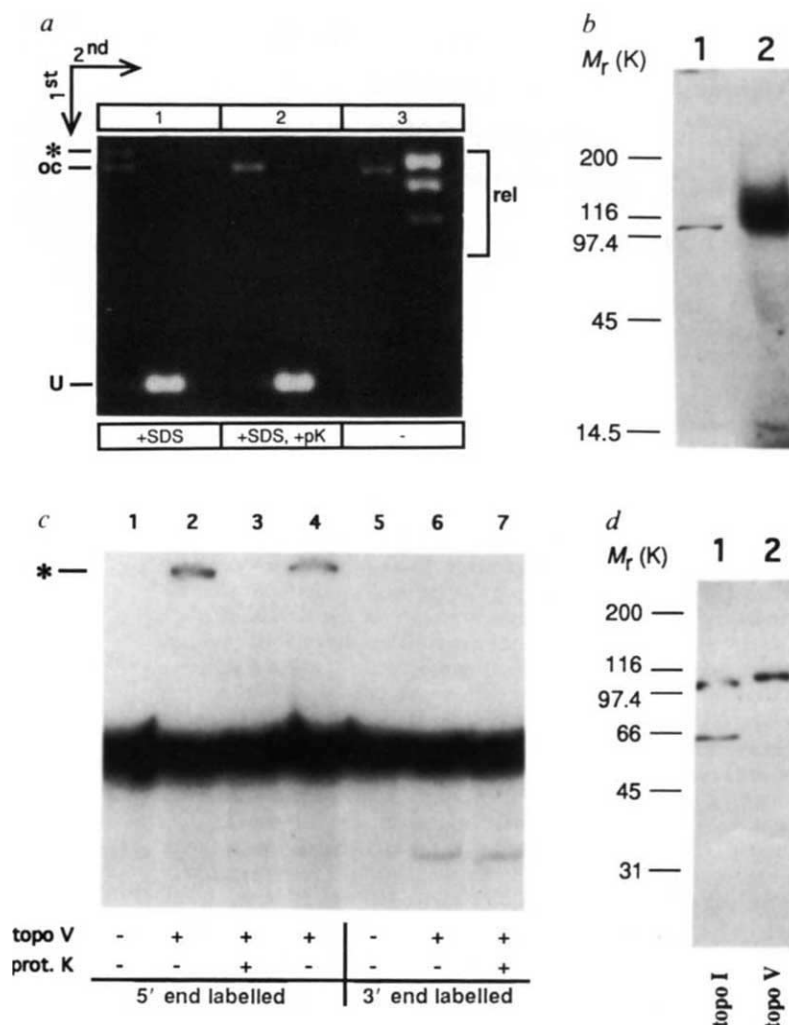
FIG. 1 Time course (a) and temperature dependence (b, c) of topoisomerase V activity.

METHODS. Batch cultures of *M. kandleri* were grown as described⁹. The purification procedure of topoisomerase V to near-homogeneity (2×10^6 U mg^{-1}) with 40% recovery followed the common procedures of cell breakage, removal of nucleic acids by polymin P, precipitation by ammonium sulphate, and column chromatography through phosphocellulose, heparin-Sepharose and gel filtration^{16,19}. The standard assay of topoisomerization activity consisted in an incubation of an aliquot with 0.2 μg supercoiled pBR322 DNA in 30 mM Tris-HCl, pH 6.7 (measured at 88 °C), 1 M potassium glutamate, 5 mM EDTA, at 88 °C for 15 min. The reaction was terminated by cooling and addition of SDS to 1%. Products were analysed by electrophoresis on a 1.5% agarose gel. Positively (+sc, lanes 1–5 of a and b) or negatively (-sc, lanes 5–10 of a and c) supercoiled pBR322 was incubated with 2 ng (a) or 20 ng (b, c) of topoisomerase V for different lengths of time (a) or at different temperatures (b, c). Chloroquine at 1.6 $\mu g\ ml^{-1}$ (a) or 25 $\mu g\ ml^{-1}$ (c) was added during gel electrophoresis. The mobility of different forms of closed circular DNA in chloroquine-containing gels is discussed in ref. 16. U, unwound DNA; oc, open circular DNA; rel, relaxed circular DNA.

† A.I.S. and S.A.K. are affiliated with the Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia, and the R. E. Kavetsky Institute of Oncology and Radiobiology, the Ukrainian Academy of Sciences, Kyiv 252127, Ukraine respectively. Correspondence should be addressed to A.I.S. at UCLA.

FIG. 2 Covalent complex formation (a–c) and recognition of topoisomerase V by anti-human topoisomerase I antibody (d).

METHODS. a, 0.2 µg relaxed pGEM3 (rel) was incubated in 30 mM Tris–HCl, pH 6.7, 300 mM NaCl, 5 mM Mg(O₂CCH₃)₂ either with 50 units *Mka* topoisomerase V (lanes 1, 2) or without it (lane 3) at 90 °C for 3 min, followed by the addition of SDS to 1% at 90 °C for 1 min. The sample in lane 2 was then digested with proteinase K (pK). Two-dimensional 1.5% agarose gel electrophoresis was in TAE buffer (40 mM Tris–acetate, 2 mM EDTA (pH 8.0)) with 0.1% SDS in the first dimension and 4 µg ml⁻¹ chloroquine in the second. The covalent topoisomerase–DNA complex is indicated by an asterisk in a and c. b, Uniformly labelled pBR322 was prepared using the QuickPrime Kit (Pharmacia) and [α -³²P]dCTP, then passed through a G-50 spin column. Labelled DNA (5 ng) was incubated with 40 ng topoisomerase V in 20 mM Tris–HCl, pH 7.0, 0.1 M NaCl at 70 °C for 3 min. The reaction was terminated by KOH^{20,21} and the bulk DNA was digested by DNase I and exonuclease III. Products were analysed by electrophoresis on 4–15% gradient SDS–polyacrylamide gels and exposed to X-ray film (lane 2). Lane 1, SDS–PAGE of topoisomerase V (silver-stained gel). c, Procedure as in a, except that the substrate was a 283-bp DNA fragment labelled at either its 5' or 3' end, incubation was at 80 °C, and EDTA was added instead of magnesium acetate for lane 4. Samples were desalted, boiled in loading buffer, run on urea–formamide 6% PAGE and exposed to X-ray film. d, Human topoisomerase I (50 ng) and *Mka* topoisomerase V (200 ng) were electrophoresed on a 7.5% SDS–polyacrylamide gel and blotted onto Immobilon P membrane. The blot was probed with rabbit anti-human topoisomerase I antibody. Immunoreactive material was detected with ¹²⁵I-labelled protein A. Human topoisomerase II, λ Int protein, *E. coli* topoisomerase I, *S. cerevisiae* topoisomerase III and antibody to *E. coli* topoisomerase I were used as negative controls (results not shown).



Topoisomerase V changes the linking number of a unique topoisomer in steps of one (data not shown), which classifies it as a type-I topoisomerase. The addition of a fivefold excess of single-stranded ΦX174 DNA to the substrate pBR322 plasmid produces no effect on the relaxation (not shown). This result, together with the ability to relax positively supercoiled DNA, means that topoisomerase V recognizes double-stranded regions of DNA.

Treatment of the reaction mixture with protein denaturant results in the covalent binding of topoisomerase V to DNA. The retardation of DNA by an attached enzyme molecule is shown in Fig. 2a, lane 1. Proteinase K digests this bound topoisomerase and yields nicked DNA circles (Fig. 2a, lane 2). The transfer of a radioactive phosphate from a uniformly labelled DNA to the protein, followed by digestion of the bulk DNA with DNases is shown in Fig. 2b, lane 2. Using denaturing PAGE, a protein–DNA covalent complex is observed when a linear 5' end-labelled DNA is used as substrate; liberated shorter DNA fragments are seen when the label is at the 3' end (Fig. 2c). Thus *M. kandleri* topoisomerase V binds to the 3' end of the broken DNA strand, as does eukaryotic topoisomerase I (ref. 17). It is interesting that the one topoisomerase V cleavage site that we have so far identified (ACACTA↓TAGAATACAAG; data not shown) resembles the preferred DNA cleavage site of several eukaryotic topoisomerase I species¹⁸. Further experiments will be needed to show whether these topoisomerases have a common motif for cleavage.

An additional line of evidence that the eukaryotic and prokaryotic enzymes may be at least partly homologous is provided by

the reaction of topoisomerase V with antibody raised against human topoisomerase I (Fig. 2d), but not with antibody against *Escherichia coli* topoisomerase I.

The activity of topoisomerase V in high salt and at high temperature could be useful in the study of DNA in these unusual conditions. Comparison of prokaryotic topoisomerase V and eukaryotic topoisomerase I will lead to a deeper understanding of the structural and biochemical relationships within this group of enzymes and will give clues about the path of evolution bordering between pro- and eukaryotes. □

Received 18 December 1992; accepted 25 May 1993.

1. Gellert, M. A. *Rev. Biochem.* **50**, 879–910 (1981).
2. Wang, J. C. A. *Rev. Biochem.* **54**, 665–697 (1985).
3. Maxwell, A. & Gellert, M. *Adv. Protein Chem.* **38**, 68–107 (1986).
4. Wang, J. C. *J. Biol. Chem.* **266**, 6659–6662 (1991).
5. Kornberg, A. & Baker, T. A. *DNA Replication* 379–401 (Freeman, New York, 1992).
6. Wang, J. C. & Liu, L. F. in *DNA Topology and its Biological Effects* (eds Cozzarelli, N. R. & Wang, J. C.) 321–340 (Cold Spring Harbor Laboratory Press, New York, 1990).
7. Wang, J. C., Caron, P. R. & Kim, R. A. *Cell* **62**, 403–406 (1990).
8. Huber, R., Kurr, M., Jannasch, H. W. & Stetter, K. O. *Nature* **342**, 833–834 (1989).
9. Kurr, M. et al. *Arch. Microbiol.* **156**, 239–247 (1991).
10. Burggraf, S., Stetter, K. O., Rouviere, P. & Woese, C. R. *Appl. Microbiol.* **14**, 346–351 (1991).
11. Rivera, M. C. & Lake, J. A. *Science* **257**, 74–76 (1992).
12. Woese, C. R., Kandler, O. & Wheelis, M. L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4576–4579 (1990).
13. Perler, F. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5577–5581 (1992).
14. Ouzounis, C. & Sander, C. *Cell* **71**, 189–190 (1992).
15. Kietzin, A. *Nucleic Acids Res.* **20**, 5389–5386 (1992).
16. Siesarev, A. I., Zaitzev, D. A., Kopylov, V. M., Stetter, K. O. & Kozyavkin, S. A. *J. Biol. Chem.* **266**, 12321–12328 (1991).
17. Champoux, J. J. in *DNA Topology and its Biological Effects* (eds Cozzarelli, N. R. & Wang, J. C.) 217–242 (Cold Spring Harbor Laboratory Press, New York, 1990).
18. Bonven, B. J., Gocke, E. & Westergaard, O. *Cell* **41**, 541–551 (1985).

19. Slesarev, A. I. *Eur. J. Biochem.* **173**, 395–399 (1988).
 20. Tse, Y.-C., Kirkegaard, K. & Wang, J. C. *J. biol. Chem.* **255**, 5560–5565 (1980).
 21. Kovalsky, O. I., Kozavkin, S. A. & Slesarev, A. I. *Nucleic Acids Res.* **18**, 2801–2805 (1990).

ACKNOWLEDGEMENTS. We thank G. Frey for *M. kandleri* cells, L. Liu for antibody against human topoisomerase I, H. Nash for λ Int protein, Y. C. Tse-Dinh for *E. coli* topoisomerase I, J. Wang for *S. cerevisiae* topoisomerase III and antibody against *E. coli* topoisomerase I, and C. Scheimer and K. Eichinger for technical assistance. The work was supported in part by funds from the Leibniz Preis (K.S.) and the NSF (J.A.L.), and by Alexander von Humboldt and Alfred P. Sloan fellowships (to A.I.S.).

CORRECTIONS

Mobile reactive centre of serpins and the control of thrombosis

R. W. Carrell, D. Li. Evans & P. E. Stein

Nature **353**, 576–578 (1991)

OUR study demonstrated the mobility of the reactive centre loop of antithrombin by its induced insertion into the A-sheet of the same molecule to give the latent L-form. We have since observed that a related process simultaneously occurs owing to the insertion of the reactive centre loop of one molecule into the A-sheet of another. We have rechecked our original data and find that, although the resulting loop-sheet polymers share similar properties with the L-form, their presence invalidates the interpretation of the CD spectra (Fig. 2 of the above letter). We have now shown that L-antithrombin is present in human plasma that has been pasteurized and we will publish separately the conditions under which it can be heat-induced from fresh plasma, in good yields free of polymers. Studies of this polymer-free L-antithrombin together with structural results by ourselves and others, confirm the conclusions in our paper as to the occurrence of the L-state, its properties and the accompanying changes in the heparin affinity of antithrombin. The new data support the overall results and conclusions of the study. □

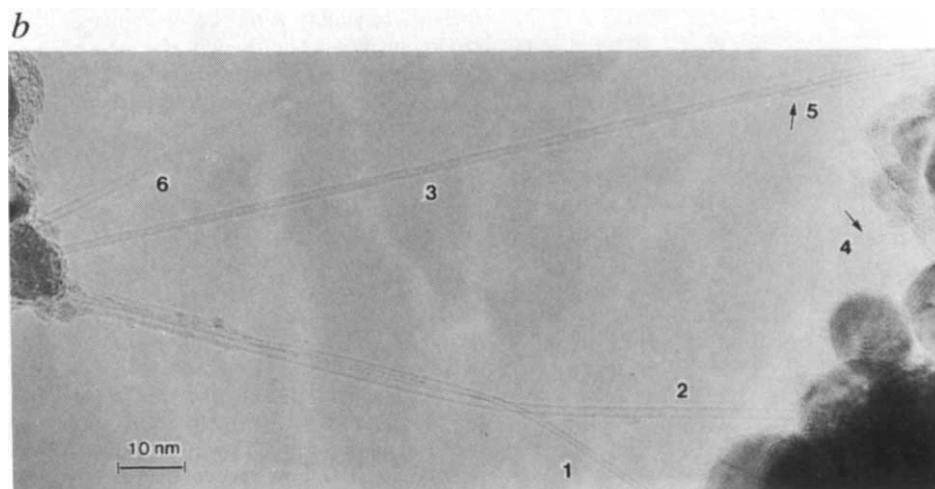
ERRATUM

Single-shell carbon nanotubes of 1-nm diameter

Sumio Iijima & Toshinari Ichihashi

Nature **363**, 603–605 (1993)

FIGURE 1b in this letter was reproduced with insufficient quality for some labelled features to be visible. The improved version on the right shows the features 4 and 5 (arrowed) that were indistinguishable previously. Also, in the sixth paragraph of the paper, “(arrow 4 in Fig. 2)” should have read “(arrow 4 in Fig. 1b)”. □



Use of evolutionary limitations of HIV-1 multidrug resistance to optimize therapy

Yung-Kang Chow, Martin S. Hirsch, Debra P. Merrill, Lawrence J. Bechtel, Joseph J. Eron, Joan C. Kaplan & Richard T. D'Aquila

Nature **361**, 650–654 (1993)

SOME of the data reported in the above letter are incorrect. See Scientific Correspondence (this issue—**364**, 679; 1993) for details. □

Oncogene *ect2* is related to regulators of small GTP-binding proteins

Toru Miki, Cheryl L. Smith, Jason E. Long, Alessandra Eva & Timothy P. Fleming

Nature **362**, 462–465 (1993)

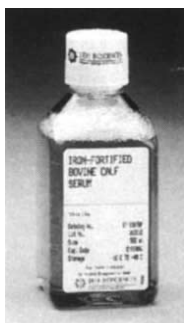
THE amino-acid numbering of Ect2 and CDC24 in Fig. 2b of this letter is incorrect. The numbers for Ect2 should be 279, 330, 382, 429 and 482 (from top to bottom). Amino acids of CDC24 are numbered according to the initial publication¹⁸ and not according to the revised sequence as indicated in the figure legend. The amino-acid numbers for CDC24 according to the revised sequence²³ should be 280, 325, 376, 418 and 470. □

Chemical science capsules

Nearly 240 exhibitors are expected to gather in Chicago next week for the 206th National Meeting of the American Chemical Society, some of whom are featured below.

WATERS Chemical Products Division of Millipore has introduced a new line of Waters Styragel HR columns for low-molecular-weight analysis of polymers (*Reader Service No. 100*). The columns are available in three solvents: THF, DMF and toluene. The series includes two extended molecular-weight-range columns (Styragel HR 4E and 5E) and five narrow molecular-weight-range columns (Styragel HR 0.5, 1, 2, 3 and 4). The extended molecular-weight-range columns are designed for applications requiring a linear molecular weight elution profile covering a broad molecular-weight range. See booth 208.

Bovine calf serum supplemented with iron is offered by JRH Biosciences as a replacement for the use of fetal bovine serum (*Reader Service No. 101*). The company says that the growth promotion capabilities of the serum are greatly enhanced by the addition of iron. A variety of cells have been successfully propagated in the serum: it has been used to support the long-term growth of fibroblast-like, epithelial-like and lymphoid populations, and has been shown to support the clonal growth of several of these cells.



Iron-fortified bovine calf serum.

The DL67 titrator from Mettler-Toledo features titration at a keystroke with two burettes and memory of the data of up to 60 samples (*Reader Service No. 102*). Due to its interfaces, extensive automation is possible with the DL67. It can be integrated into



Mettler-Toledo's DL67 titrator.

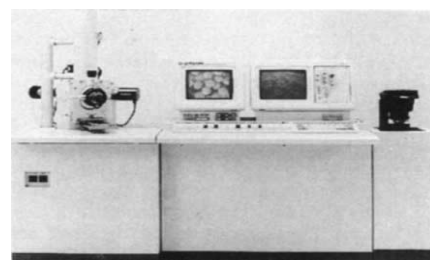
laboratory information management systems. Remote control by a computer, the attachment of a balance, a robot or different types of printer are also possible. In accordance with good laboratory practice guidelines, the DL67 records all determinations with date, time, user, instrument, titration stand, methods, sample data and result. With the appropriate accessories, it is also suitable for water content determinations by the Karl Fischer method or other voltammetric analyses. See booth 811.

Small-world solutions

By combining the atomic force microscope (AFM) with the high-resolution inverted optical microscope, TopoMetrix has developed the new Explorer Life Sciences AFM which is designed specifically for optimized atomic force microscopy use in cell biology, biophysics and molecular biology applications (*Reader Service No. 103*). The Explorer Life Science AFM integrates high-magnification imaging (up to $\times 10^9$) and three-dimensional, quantitative measurement capabilities of an AFM with any one of several commercially available inverted optical microscopes, including Zeiss, Nikon and Olympus models. The high magnification of the inverted microscope is designed to complement the imaging capabilities of the atomic force microscope by providing an independent monitor of the tip-sample interaction. The inverted optical microscope is also used in precision positioning of the AFM tip and sample. The system incorporates a new scanner design, which operates in liquid without the use of separate fluid cell hardware. Due to the liquid environment, TopoMetrix says that meniscus forces are eliminated, and imaging can be performed at much lower forces than when performed in ambient air. The microscope's One-Pass Imaging capability permits the simultaneous acquisition of up to eight channels of data in a single scan. New preparation techniques permit samples to be firmly anchored while being scanned under liquids. It is these technologies, combined with the liquid scanning techniques, that allow the instrument to be used on living systems under physiological conditions. Several traditional AFM imaging modes are offered, including lateral force, modulated force and noncontact imaging.

Universal is the new scanning electron microscope from Hitachi Scientific Instruments that is designed to produce high resolution at low and high vacuums (*Reader Service No. 104*). Low vacuum microscopy can be important in situations where sam-

ples are 'wet' with a high oil/water content. These include plant and animal specimens, foods, pharmaceuticals and paints, which are best observed in their 'natural' state (that is, without pretreatment) at low vacuums. Hitachi says that the same is true for semiconductor materials. As samples do not need to be pretreated, observation is a simple procedure and occurs without the introduction of artefactual features that can arise from the treatment process. Universal has a stated magnification of up to $\times 200,000$ and a guaranteed resolution of 6 nm at low vacuum. The microscope has the usual range of add-on features, including elemental analysis (EDX and WDX are both available



Hitachi's Universal SEM — works well at low and high vacuums.

at low or high vacuums). The instrument also features two $1,024 \times 1,024$ pixel frame stores and two 512×512 pixel frame stores. Recorded images may be displayed on a four-quadrant split screen for ease of comparison. A wide range of processing functions are provided to improve the quality of the image; these include a recursive filter, image integration, polarity reverse, and digitization. Images can be stored onto an optical archiving system, output on a thermal printer or recorded by camera. The specimen chamber accommodates samples up to 150 mm in diameter and a choice of three specimen stages is available. External computer control and Cryo operation are optional. See booth 520.

Molecular biology

Sigma's 1993-1994 peptides and amino acids catalogue provides the user with a selection of over 4,000 peptide-related products (*Reader Service No. 105*). New products range from reagents for protein analysis and peptide synthesis to highly



Sigma's peptide/ amino acid sampler.

characterized bioactive peptides, substrates, inhibitors, amino acids and polyamino acids. Highlighted products include new β -amyloid fragments, neuropeptide Y analogues and HIV protease substrates. In addition, the cleavage and deprotection of peptide resins is now designed to be safer and easier with the introduction of a new peptide deprotection kit that replaces the need for liquid hydrogen fluoride in solid-phase peptide synthesis. See Sigma Aldrich in booth 745.

The SpinBind **mini-prep system** from FMC BioProducts is designed to offer users a fast and efficient way of obtaining high-quality plasmid DNA from bacterial culture (*Reader Service No. 106*). Using this system, there are no resins to resuspend, no phenol extractions or ethanol precipitations, and no RNase. Plasmid DNA purified with this system can be used immediately for restriction digestion, sequencing and cloning. The complete system contains all the necessary reagents; no syringes or dedicated equipment are required. See the FMC Lithium Division in booth 234.

Transform *Escherichia coli* conveniently, economically and efficiently with the new **Pulser transformation system** from Bio-Rad, details of which can be found in a bulletin number 1734 which is available from the company (*Reader Service No. 107*). The Pulser is designed to deliver the optimal 5-ms time constant to *E. coli* samples of high resistance. The user does not need to choose a parallel resistor or capacitor, as the decision has already been made. The instrument can accommodate either 0.2 cm or 0.1 cm electrode-gap cuvettes, and delivers the optimized field strength to either cuvette with easily accessed pre-programmed voltage settings. The instrument is not limited to these two voltage settings, however; a choice of voltages between 200 V and 2,500 V is available. To ensure reproducibility, both the time constant and the voltage delivered may be verified for every pulse. Seven strains of Electro-Competent cells are offered for use with the *E. coli* Pulser, including popular cloning and mutagenesis strains (some with blue/white selection properties). Bio-Rad says that use of these cells reduces the preparation time to 5 min. See Bio-Rad's Digilab Division in booth 240.

■ Spectroscopy

BenchTOF is the new space-saving **time-of-flight mass spectrometer** from Bruker that is designed as a dedicated instrument for the analysis of biologically important molecules (*Reader Service No. 108*). Despite its small size, the BenchTOF still offers users ± 35 kV operation in both linear and/or reflection modes to provide simultaneously high resolution and high sensitivity. Standard options include both a unique *x,y* position multiple sample probe for automated analysis and high-definition observation

optics for sample manipulation. The system is computer controlled and operated with a user-friendly, graphical mouse-driven interface. Bruker is also making available peptide and protein databases for rapid online library searches to identify unknown samples or impurities. Stated applications of the BenchTOF include direct analysis of peptides and proteins, enzymatic digests, carbohydrates and glycoproteins, and synthetic polymers. See booth 916.

Details of the new Model 3110 **atomic absorption spectrometer** from Perkin-Elmer, which features built-in control of the flame, graphite furnace and flow injection devices, can be found in a new 12-page colour brochure available from the company (*Reader Service No. 109*). The brochure



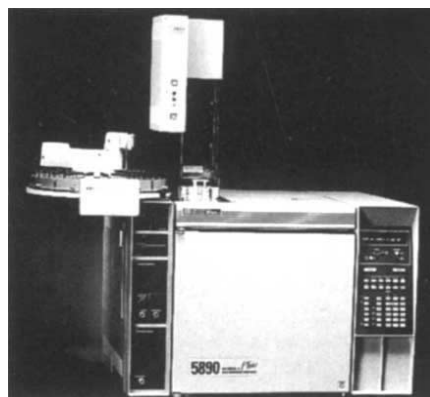
Model 3110.

highlights some of the design features of the Model 3110, including a high-dispersion monochromator (with dual-blazed grating) for maximum light throughput and improved performance, and covered quartz-coated reflecting optics for high energy efficiency over the entire atomic absorption wavelength range. It also describes the spectrometer's burner control system, which provides full operational safety, in addition to a built-in microcomputer. The enhanced data system option provides advanced data handling, reporting and control of accessories. Additional accessories for the Model 3110 include the HGA graphite furnace and the flow injection analysis systems. See booth 326.

Bruker says that researchers working in the life sciences do not need to become expert mass spectroscopists to use its new ProteinTOF **matrix-assisted laser desorption time-of-flight mass spectrometer** (MALD-TOF-MS) (*Reader Service No. 110*). As the system is run entirely via an Apple Macintosh computer, protein chemists are provided with an intuitive, user-friendly graphical software environment. As an option, ProteinTOF is also available with a PeptideSearch software tool for comparing acquired MALD-TOF-MS data with 'computer digests' of various EMBL and NIH databases, for the rapid identification of known proteins and nucleic acids. Bruker also announces the availability of FAST (fragment analysis and structural TOF-MS) — a new accessory that provides fragmentation and structural information directly from a Bruker REFLEX mass spectrometer for information on structure, sequence and post-translational modifications.

■ Gas chromatography

The new HP 5890 Series II Plus **gas chromatograph** from Hewlett-Packard features six-channel electronic pressure (EPC) control (*Reader Service No. 111*). The addition of four auxiliary EPC channels to the two existing EPC channels is designed



HP 5890 Series II gas chromatograph.

to provide the instrument with precise digital control of all gas chromatograph gases for both inlets and detectors. This arrangement provides users with the same level of digital programming capability and ease of controlling gas flows as they may be accustomed to for control of the column oven. Hewlett-Packard says this feature enhances performance and precision of the gas chromatograph, increases productivity and reduces operating costs. Two EPC channels are dedicated to inlets and operate in several modes: constant pressure, constant flow in capillary columns, pressure programming and vacuum compensation for use with mass spectrometers. The four auxiliary EPC channels can optimize performance for a variety of applications: they provide control for carrier gases, reagent gases and detector makeup and fuel gases. They also control flows for sample introduction devices such as the HP 7694 headspace sampler and purge and trap units. Combining the instrument with the HP 7673 automatic injection/sampler system and the HP 3365 Series II ChemStation makes the gas chromatograph a fully automated research and quality control tool. See booth 601.

Varian has introduced a new version of its Star Workstation **gas chromatography software**, which is designed to enhance laboratory productivity through an interactive data manager and a new software automation tool called AutoLink (*Reader Service No. 112*). The curve manager allows chromatographers to review calibration data, both graphically and statistically, collected over nine different calibration ranges. Varian says that this makes it easier and faster to construct accurate calibration curves, avoiding mistakes that can lead to costly repetitions of analyses. The Star Workstation verification mode further enhances productivity by automatically checking calibration curves



Varian's gas chromatography workstation.

for accuracy once they are constructed. This system check reduces the need for frequent recalibration of the instrument, while still providing full documentation for regulatory compliance. The new AutoLink feature utilizes dynamic data exchange to link automatically the instrument's system software to specialized analysis software. Before, such programs had to be initiated manually after an analysis. The gas chromatography workstation supports up to four gas chromatographs and can collect data from up to eight detectors, doubling its previous workload capacity. See booth 947.

■ UV/visible range

The PC U-2000 is the new PC-controlled, double beam **ultraviolet/visible spectrophotometer** from Hitachi Scientific Instruments (*Reader Service No. 113*). Based on the U-2000 UV/visible spectrophotometer, the PC U-2000 is available in two versions: the PC U-2000DOS and PC U-2000WIN. Each version comes with a 33-MHz 386 processor IBM-compatible personal computer, with a 14-inch VGA colour monitor, 2-MB RAM, 42 MB hard disk, 3.5-inch floppy disk drive, keyboard and mouse. The PC U-2000DOS features MS-DOS version 5.0 and data manager software, and is designed to control the U-2000 and acquire wavelength and time-based spectral information for display and post-run processing. The multi-window display and real-time scale expansion allow experimental results to be verified in a real-time environment. Routine spectroscopic methods can be set up, stored and retrieved, while data can be stored or exported directly to Spectra Calc software. For users who prefer the Windows interface, the PC U-2000WIN offers the benefits of Windows version 3.1. The software includes a special data retrieval package, which supports photometry, wavelength scan, time scan and multiple wavelength operations, allows data to be manipulated in the Windows environment, and is compatible with the U-2000 autosampler.

A new six-page brochure from Perkin-Elmer describes the company's new Lambda 11 **ultraviolet/visible spectrometer**, which is intended as a low-cost scanning spectrometer for use in routine analyses and educational applications (*Reader Service No. 114*).

With this instrument, the methods can be configured according to the five primary laboratory tasks: these include spectral scanning, absorbance-versus-time analysis, wavelength programming with up to eight wavelengths, and concentration analysis. Various accessories are available for the Lambda 11.

■ Databases and software

Synopsis Scientific Systems, a supplier of computerized scientific information products and services, announces the release of version 2.0 of its chemical reaction **database of protective group chemistry** (*Reader Service No. 115*). Providing a thorough and systematic coverage of the literature on protective group chemistry, the database includes up-to-date methods for protecting and deprotecting specific functional groups, including important information on regio- and chemoselective control. The database also incorporates cross-referenced data on tolerated groups and stability and lability of the protected groups — the kind of information that a chemist needs when selecting protection and deprotection reactions to be used during a synthesis. Version 2.0 of the database contains 9,000 reactions abstracted from over 3,000 primary literature references. It contains a number of enhancements suggested by users, including standardized keyword conventions and refined selection criteria. The company says that this version provides additional stability data and wider coverage of some key areas of protecting group chemistry, such as peptides, carbohydrates and nucleosides. See booth 414.

The first major upgrade to Autodesk's scientific **software for molecular simulation**, HyperChem, has been launched by the company's sole UK distributor Techex (*Reader Service No. 116*). HyperChem operates under the standard Microsoft Windows graphical user interface and enables chemists and other scientists to build, manipulate and study three-dimensional molecular structures on a personal computer, and to use the data and graphical output with other Windows-based software programs. The new version of HyperChem Release 3 features extended visualization, computation and data integration features, making it a useful tool for data analysis and in the preparation of reports and presentations. New and enhanced features in HyperChem Release 3 are designed to expand its utility in five main areas: predicting and analysing infrared, ultraviolet and visible spectra; visualizing protein structures using ribbons to display backbones; building and analysing molecules containing transition metals; transferring data to and from other applications programs commonly used in research; and for reading and writing to the file formats of industry-standard databases for molecular modelling and computational chemistry with the ability to add user-definable

file formats. These new features are designed to expand the program's utility in a variety of research disciplines, including the pharmaceutical industry, medical biotechnology, synthetic organic chemistry and materials science. Upgrades for users of HyperChem Release 2 are available. See booth 434.

Sybyl/Molcad, an interactive graphics module for the **display and manipulation of molecular structures** and the visualization of related data, is now available from Tripos, as the result of an exclusive distribution agreement between Tripos and Professor Jürgen Brickmann of the Technische Hochschule Darmstadt in Germany (*Reader Service No. 117*). Tripos says that the Molcad module provides unique and essential visualization of molecule characteristics, such as Connolly surfaces, electron density surfaces, and various properties. The properties currently available include lipophilic potential, electrostatic potential, hydrogen bonding and local curvature. Molcad calculates values for these properties and maps them, via colour-coding, onto Connolly and electron density surfaces. It is available on the Silicon Graphics IRIS series and soon on the Evans and Sutherland ESV. The company is also making available Leapfrog *de novo* design software, which can be used to propose new compounds based on enhanced properties. See booth 710. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENT

Plant and Animal Tissue

HOMOGENIZER and EMULSIFIER

- Disrupts and disperses small samples (1/2-100 ml)
- Both mechanical and ultrasonic action
- 5,000-30,000 rpm
- Built-in speed control
- Easy to clean
- Best price (\$485 complete)



The TISSUE-TEAROR™

Manufactured By:
BIOSPEC PRODUCTS • PH/FAX (918) 336-3363
P.O. BOX 722 • BARTLESVILLE, OKLAHOMA 74005

Reader Service No. 15

Science research and philanthropy

Differences in culture, industrial practices and government support mean that the role of philanthropic foundations and charities in supporting research differs in the US, Europe and Japan.

Government funding dominant

Washington. Despite the US government's dominant role in funding science, there is no doubt that political and economic shifts threaten the long-term stability of research and the job security of scientists. And, although private industry has increased research funding during the past decade to levels now approaching those of federal expenditures, thousands of scientists in the private sector are also finding their jobs jeopardized by recessionary cutbacks. As long as politics and the economy continue to reshape the roles that government and industry play in funding science, there remains one consistent, albeit limited, alternative source of support: private philanthropies. With contributions of \$680 million in 1992 — or about 3 per cent of the \$25.5 billion spent on nonmilitary scientific research that year — private trusts, endowments and foundations helped to fund an estimated 32,000 research positions nationwide.

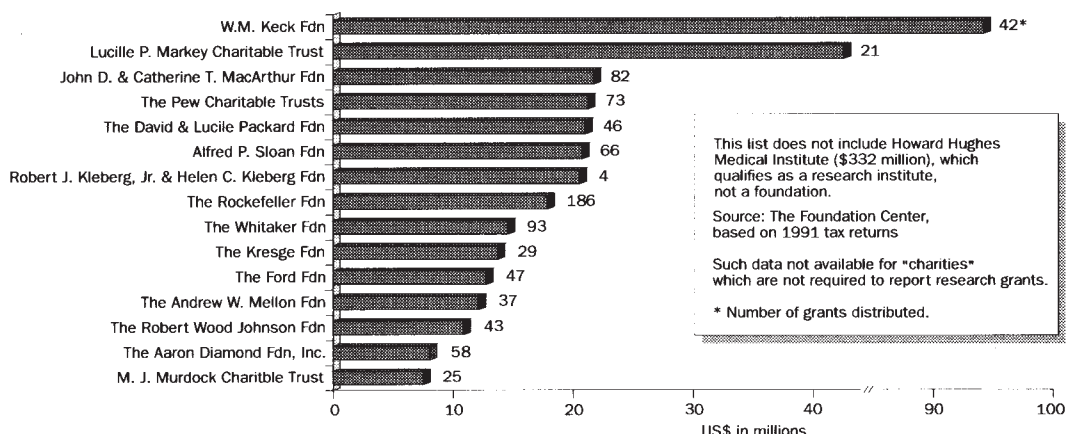
Philanthropies' most significant contribution often comes from their support of young scientists. To counter trends away from careers in science, which many fear may lead to a shortage of researchers in the new century, foundations have stepped-up funding for education, particularly predoctoral and postdoctoral fellowships. Many have also redirected their money towards young researchers who have completed their training but are unable to win funds from the hugely competitive federal agencies, including the National Institutes of Health (NIH) and the National Science Foundation (NSF). According to the US Bureau of Labor Statistics, the number of unemployed natural scientists has increased during the past three years by two-thirds, from 15,000 to 25,000. Furthermore, aerospace and defence companies, drug firms and big research universities are all facing pressure to limit funding for basic science. The trend appears unlikely to change in the near future.

"Funds have dried up for younger scientists," says Robert Glaser, director of medical science and trustee for the Lucille P. Markey Charitable Trust. "We're now seeing foundations coming to the aid of young investigators who have new ideas and need help substantiating them for the national scene."

With total assets of \$6.9 billion, the Howard Hughes Medical Institute (HHMI) is by far the largest philanthropy supporting research. In 1992, the institute spent \$51.5 million on 296 predoctoral fellowships in the biological sciences, 100

considerable impact on the science community. For years, the Pew, Packard, Pfizer, Searle and Sloan foundations have been major players in financing the education and research of young biomedical scientists. Since 1984, the Lucille P. Markey Charitable Trust has topped the list, with expenditures averaging \$51 million a year. However, as stipulated by the will of its founder, the trust is scheduled to spend itself out of existence by 1997 when, coincidentally, a final instalment of a \$400-million windfall will thrust the Burroughs Wellcome Fund into the position of the nation's top private biomedical grantmaker after Hughes. Although the Burroughs Wellcome's annual giving of

Top US foundations in science or medical research, 1991



research training fellowships for medical students and 72 postdoctoral fellowships for physicians who have completed at least two years of postgraduate clinical training. As for its funding of more established scientists, Hughes often supports investigators throughout their careers. Last year, it spent \$281 million employing 224 investigators in the fields of cell biology, structural biology, genetics, immunology and neuroscience. The institute selects leading medical investigators and hires them as employees at their home institutions, where HHMI pays their salaries plus the costs of staff, technical assistance, equipment, supplies and, in many cases, new or renovated laboratories.

It is a testament to the Hughes Institute's wealth that each year it gives more to scientific research and education than the 100 largest philanthropic foundations put together. This is not to say, however, that smaller foundations do not have

\$22 million will be considerably less than Markey's \$55 million, the fund is designed to last in perpetuity. Trust officials say they will spend the next year deciding which areas of research they will support.

When looking for grants, it is important that researchers understand the differences between philanthropies and charities. Scientists seeking to conduct research on certain diseases may do best to apply to charities focusing specifically on those issues. For example, this year the American Heart Association will give out \$105 million in research grants to scientists studying various aspects of the cardiovascular system. Other top charities supporting disease-specific research include the American Cancer Society (\$94 million), the March of Dimes Birth Defects Foundation (\$14 million) and the American Lung Association (\$7 million). Unlike charities, private philanthropies don't usually dedicate themselves to fund-

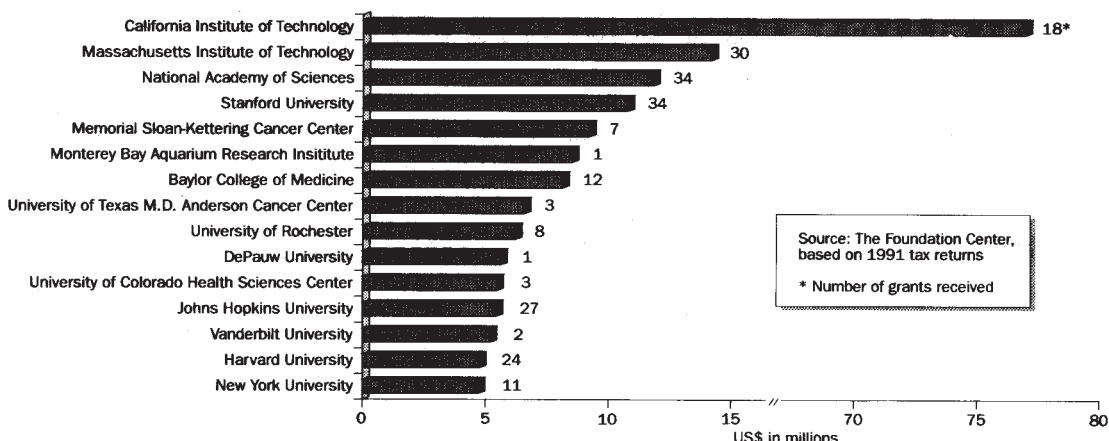
ing research on specific diseases. Instead, most try to maximize the value of their funds by concentrating on new fields that they feel deserve particular attention for limited periods of time. Currently, the hottest among these are genetic research, neuroscience, molecular biology, immunology, AIDS, breast cancer and women's health, ageing and preventive medicine.

Because most private philanthropies typically do not have enough money to be the primary resource of research funding in any given discipline, many hope that their involvement can help nudge, or raise awareness about particular fields. The Aaron Diamond Foundation, for example, was among the first philanthropies to fund AIDS research in the mid-1980s. Several foundations then followed Diamond's example in supporting AIDS research, now a major funding area among private philanthropies. Rather than focusing on initiating work in new fields, some foundations choose instead to rally behind projects that seek to change the way science is done in existing disciplines. The MacArthur and Dana foundations, for example, have funded innovative work in the neurosciences resulting in the development of multidisciplinary approaches to brain research and maintaining mental health.

Such specialized funding is a mark of how the role of foundations in the support of science has changed over the years. From the late 1800s to the 1940s, private foundations were the major independent source of funding for all biomedical research. In 1940, before the rise of the NIH, foundations provided some 27 per cent of the total \$45 million spent on health-related research. By 1986, total funding had increased many fold, to \$14.8 billion, while foundations, voluntary health associations and the private academic sector together contributed only about 5 per cent of the total. Although some of the old family philanthropies, like the Rockefeller and MacArthur foundations, are still active in the sciences, many, such as the Carnegie (family) Foundation, retreated after the growth of NIH with the idea that foundations could best use their resources elsewhere. There are now 32,000 private grant-making foundations in the United States giving away \$4.7 billion annually; only an estimated 2 per cent of those, or a total of 700 foundations, now support scientific research.

"We've seen foundations turn away from research toward support for social

Top US grant recipients in science and medical research, 1991



and community issues. The perception has been that science is getting from other sources the kind of funding it needs," said Martha Peck, executive director of the Burroughs Wellcome Fund, which is the

newest big player funding biomedical research. "I think it's time to change that perception, to get philanthropies back involved with supporting breakthroughs in American science." **Susan Greene**

Charities taking the strain

London. Faced with a combination of reduced government funding — the main source of support for basic science in the post-war period — and the continuing reluctance of industry to invest directly in long-term fundamental research projects, European scientists are turning increasingly to private foundations and charities for support.

The prominence of such organizations varies from one country to another. France, for example, which has enjoyed a long tradition of generous state backing for public activities, has relatively few. In contrast, private organizations play a larger role in Germany. Many of these are funded out of the profits of successful industrial enterprises; the largest, for example, is the Volkswagen Foundation, which has given out research grants totalling DM3.9 billion (\$2.3 million) over the past thirty years — DM158.8 million in the past year alone.

These figures are still relatively low, however, compared to the German government's funding of science. In other, smaller countries, foundations linked to multinational corporations — such as the Carlsberg Foundation in Denmark, financed by the brewing company of the same name (see *Nature* 360, 519; 1992) — are able to play a significant part in the national research effort as a direct result of their size relative to public research agencies.

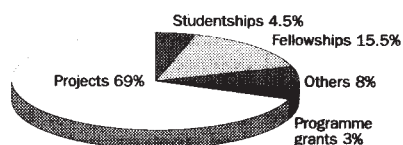
But the country in which foundations and charities have come to play the most prominent role is Britain. Statistics alone reveal much of the story; according to figures published last month by the Uni-

versity Statistical Record, the income of British universities from such bodies rose by 110 per cent between the years 1986–87 and 1990–91, more than twice the increase in their income from the research councils over this period, and considerably more than the 69 per cent increase from industry.

These bodies now provide almost one quarter of the research income from universities — up from 19 per cent in the mid-1980s. But aggregated statistics do not show the whole picture. In some research fields, charities now play the dominant role. Last year, for example, the 73 charities which belong to the Association of Medical Research Charities spent a total of £266 million on research, considerably more than the Medical Research Council's expenditure of £200 million. And, having edged ahead for the first time in 1990–91, they are increasing their lead.

Many of Britain's most prominent charities have, in the past, been based on the philanthropy of successful industrialists. The Nuffield Foundation, for example, one of the biggest private sponsors of research in the period after the Second World War, was set up originally through

AMRC charities: types of grant



Total Number of Grants, 1990 and 1991 = 2,910
Not included: grants awarded to charities' own institutions

a £10 million bequest from the will of the motor manufacturer William Morris, later Lord Nuffield.

As Britain's traditional industrial base slipped into decline, so too did the fortunes of many such bodies, particularly those whose capital remained invested in the company from which they had grown. The collapse of the British motor industry in the late 1960s and 1970s, for example, had a devastating effect on the Nuffield Foundation, from which it has only recently recovered.

Many of the old giants have been replaced by those financed from newer fortunes. A major benefactor of science in recent years, for example, has been the Sainsbury family, the owners of one of Britain's largest and most successful supermarket chains. The family has now become an important source of funding for science and technology projects in schools through the Gatsby Foundation.

Perhaps the most spectacular growth has been that of the Wellcome Trust. Created by the US born pharmaceutical manufacturer Henry Wellcome, the trust was left all the shares in his company, the Wellcome Foundation, when he died in 1936. The sale of a quarter of these shares in 1986, followed by that of a further 30 per cent last year — both during a boom period in the fortunes of pharmaceutical companies in general and Wellcome in particular — have given the trust a capital investment fund of £5.4 billion (\$8.0 billion), even larger than that of the Howard Hughes Medical Institute in the United States.

As a result of its new investment strategy, Wellcome will be in a position to spend about £200 million in support of biomedical research next year, roughly the same as the whole research budget of the MRC. Much of the new money is being used to help create centres of excellence in fields such as molecular science, genome research and cell biology. It has also given the trust the ability to take initiatives (such as offering higher-than-average salaries to particular individuals in order to entice them to Britain from posts in other countries) which government-funded agencies such as the MRC would find difficult to implement.

The past decade has also seen growth in Britain's many disease-specific charities. The reasons for their success are complex; one suggestion is that these charities, over 70 per cent of whose income comes from legacies, benefit from those who have been treated under the National Health Service, and seek to make a personal contribution to medical research in return (in contrast to countries such as the US where medical care is paid for directly by the individual).

Whatever the reasons, bodies such as the Imperial Cancer Research Fund (with an annual research budget of £50 million),

THIRTY LARGEST RESEARCH FUNDING UK CHARITIES

	Expenditure (£ thousand)
Wellcome Trust	77,231
Imperial Cancer Research Fund	49,932
Cancer Research Campaign	44,199
British Heart Foundation	20,944
Arthritis and Rheumatism Council for Research	11,800
Leukaemia Research Fund	9,053
Yorkshire Cancer Research Campaign	4,845
Multiple Sclerosis Society	4,700
Ludwig Institute For Cancer Research	4,490
Action Research	3,500
British Diabetic Association	2,632
Muscular Dystrophy Group	1,931
Sir Jules Thorn Charitable Trust	1,874
National Asthma Campaign	1,829
Cystic Fibrosis Research Trust	1,644
Marie Curie Cancer Care	1,490
Ciba Foundation	1,381
North of England Cancer Research Campaign	1,319
William Harvey Research Institute	1,162
Birthright	1,100
Parkinson's Disease Society	1,093
British Lung Foundation	1,044
National Kidney Research Fund	1,010
Stroke Association	959
Wessex Medical Trust	926
Tenovus Cancer Fund	925
Motor Neurone Disease Association	900
Mental Health Foundation	835
Leverhulme Trust	824
Lister Institute of Preventive Medicine	762

the Cancer Research Campaign (£44 million) and the British Heart Foundation (£21 million) each play a role comparable in importance to that of public research bodies elsewhere, such as the individual National Institutes of Health in the United States.

In the past, when budgets were smaller, the main function of medical charities was often merely to top up the research efforts of government funded agencies, in particular the MRC. Over the past decade, however, as their size and impact has grown, so too has an awareness of the need to take a more strategic approach to their funding of research in universities and medical schools. "All of a sudden, these bodies have become mainstream funders, having to think about their research programmes and support for research fellows in a big-time league," says David Katz of University College London's School of Medicine, who is chairman of the Arthritis and Rheumatism Council's Fellowship Committee. "Many now realize that, unless they set up a rigorous review and monitoring structure, they are going to run into trouble."

One result, particularly as the demand for research grants increases with the drop-off in public funding for medical research, is closer scrutiny of grant proposals. The AMRC, for example, set up in 1987 to represent the common interests of medical research charities in their negotiations with government, insists that all its members apply peer review to grant applications. A second result has been growing concern to address the career prospects

of the scientists they support. Previously, individuals receiving postdoctoral fellowships could, for example, optimistically expect appointment to a university post when the fellowship support came to an end.

But universities are no longer expanding. As a result, science departments are becoming clogged up with high-quality scientists employed on a succession of short-term contracts, with little long-term job security and virtually no possibility of career planning.

In these circumstances, several medical charities are now developing schemes not only for providing longer-term support, but also for implementing a parallel career structure for the university-based researchers that they support, independently of the career opportunities offered by the university itself.

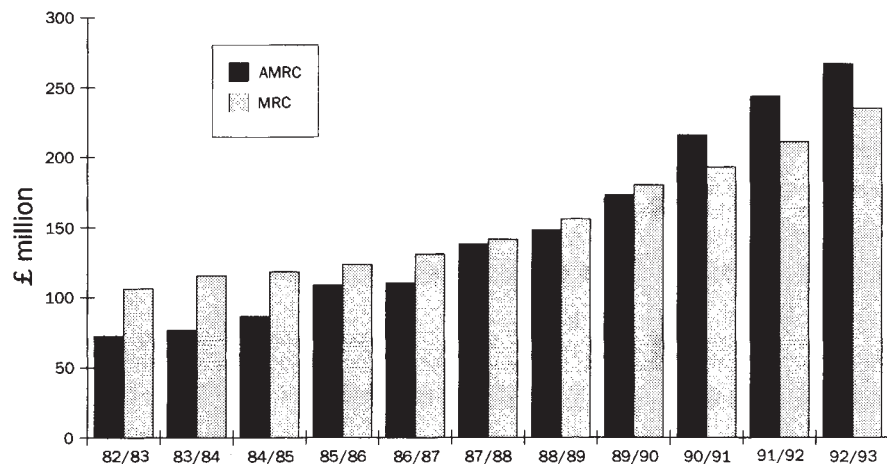
Under Katz's guidance, for example, the ARC introduced last year a new system for financing university researchers in three grades: clinical research fellows, postdoctoral fellows and senior fellowships. Each grade has a carefully-tailored selection process, involving both British and international selectors. And the awards available at each grade are linked in a way designed to provide career stability for non-clinical researchers, even if these are merely good research team members, and not outstanding scientists in their own right. "It was an essential step for the ARC to take; otherwise its research funding would have become increasingly chaotic," says Katz. "The programme is going extremely successfully in a very difficult climate."

The most significant move in this direction is that currently under discussion by the Wellcome Trust. During the 1980s, the foundation's own experience reflected the growing problem of scientists on short-term contracts; the proportion of its research money going on short-term projects increased from 22 to 40 per cent over the decade. "We decided that it could not be right, with our greatly increased income, to go on funding like this," says Bridget Ogilvie, the director of the trust. "We now want to fund more careers in the long term."

As a result, Wellcome is in the process of developing plans for a new 'rolling grant' scheme under which, for example, an individual scientist could be supported — subject to periodic review — on a succession of four-year contracts as a research fellow. After that, providing the reviews remain satisfactory, the individual would be financed through to retirement as an "advanced research fellow".

To cater for the needs of those who demonstrate an aptitude for team leadership, research fellows would, in their mid-30s, also be able to compete for senior basic biomedical fellowships. These in turn would lead, for successful

AMRC and MRC expenditure, 1982-93



candidates, to principal research fellowships, also available up to retirement.

In addition to providing a more structured career ladder for successful research scientists, Wellcome is also thinking about providing counselling facilities to those for whom it is decided not to renew research fellowships. One frequent complaint is that, at present, scientists seldom receive any guidance about their future career prospects when a project grant comes to an end.

Wellcome's proposals, which are still under discussion between the trust and the scientific community, have met with an enthusiastic reception from groups such as the Association of Researchers in Medicine and Science, which has been complaining vociferously in recent years about the difficulties faced by the increasing number of scientists working on short-term contracts.

They have also been welcomed by the government. In its recent white paper (policy document) on science, the government continues to argue that responsibility for providing stable careers rests with universities and research institutions. However it does say that it recognizes that the flexibility of universities is currently limited by the terms under which it can employ research staff; and praises the trust for its attempts to improve the career prospects of scientists.

But Wellcome itself is also wary of being taken for granted by the government, and of the danger that politicians may use the growing influence of the research charities to justify cutting back on its own support for medical research (for example, by requiring them to pay for an increasing proportion of the overhead costs of university science departments).

Diana Garnham, general secretary of the Association of Medical Research Charities, points out that charitable funding tends to be short-term, since the income of most charities varies from one year to the next. "That is why most of it tends to be project oriented; but it also

makes it essential that the government continues to underpin the longer-term needs of university departments." Medical research in the UK has been strong precisely because of the multiplicity of sources of funding, says Garnham; the effectiveness of charity money has in the past only been possible because of the government's support for the science base. "Our strength has grown out of the dual support system," under which grants to universities cover basic overheads, and research councils and charities pay the additional costs of research. "That is why we are going around telling the government it should not change the way it does things."

Whitehall has clearly been listening to these views. Proposals to transfer responsibility for university overheads away from the Department of Education (a move which would have had substantial impact on the ability of charities and foundations to fund university research projects) was rejected in the white paper, and for the time being the dual support system remains in place.

But the charities themselves realize that the boundaries are changing. And that the new environment in which they are operating in the 1990s imposes on them the need for far greater social accountability for the way in which their research funds are used than in the past. The challenge they face now is how to meet their growing responsibility for the health of the research community in the most effective way, even if this means not necessarily pandering to the interests of the most vociferous (or most wealthy) interest groups.

Increased peer review (breaking up the traditional old-boy networks of grant allocation), a greater sense of overall research strategy, and an enhanced attention to the desire of scientists for the stable career structure which universities on their own are no longer able to provide, are three steps in this direction. Others will inevitably follow.

David Dickson

A new trend for Japan

Tokyo. Charity and philanthropy in the Western sense, which are so much tied to Christian values, are alien to Japan. Many research foundations do exist but they tend to be small and are often established because of personal links between powerful researchers and companies. There is no equivalent of large research foundations supported by rich individuals like Howard Hughes (of the United States) or charitable organizations like the Imperial Cancer Research Fund and Cancer Research Campaign of the United Kingdom.

Instead, there is a myriad of little foundations which have narrowly defined purposes, tend to be very domestic in outlook, and are wrapped in the bureaucratic control of the government. Every foundation in Japan has to be linked to a government ministry or agency, and although there is no legal requirement that retired ministry officials should get a job in the foundation, it is an open secret that if a retiring ministry official is offered a job (a process called *Amakudari*, or 'coming down from heaven') permission to establish the foundation tends to proceed much more smoothly. But the ex-bureaucrat then tends to limit the scope of the foundation to suit the wishes of his former ministry.

This bureaucratic control and the lack of the Christian sense of giving for the general benefit of mankind has left Japan starved of the powerful research foundations found in the West. But there is a gradual change as more and more Western companies with Western values set up shop in Japan.

Ciba Geigy, the Swiss pharmaceutical multinational, was one of the first foreign companies in Japan to establish a research foundation — the Ciba-Geigy Foundation (Japan) for Promotion of Science. Unlike the foundations of Japanese pharmaceutical companies, almost exclusively designed to support Japanese scientists, the Ciba-Geigy offers a handful of fellowships for young European scientists to come to Japan (see *Nature* 361, 762; 1993).

Japanese companies are beginning to follow this example. For example, Yamanoichi Pharmaceutical is planning to reorganize its foundation for research on metabolic disorders this year to establish an international section to fund non-Japanese researchers, says Teruhisa Noguchi, executive vice president for research and development in the company. But Japan still has a long way to go before its foundations provide a significant source of employment for Japanese, let alone foreign, scientists.

David Swinbanks